

## SEQUENCE LISTING



<110> Palmer, Michelle A.J.  
 Gee, Melissa  
 Tillotson, Bonnie  
 Chang, Xiao-Jia

<120> Systems for Sensitive Detection of G-Protein Coupled Receptor and Orphan Receptor Function Using Reporter Enzyme Mutant Complementation

<130> 4085-235-27 CIP

<140> US 09/759,152  
 <141> 2001-01-16

<150> US 09/654,499  
 <151> 2000-09-01

<150> US 60/180,669  
 <151> 2000-02-07

<160> 9

<170> PatentIn version 3.0

<210> 1  
 <211> 6700  
 <212> DNA  
 <213> Unknown

<220>  
 <223> pICAST ALC.

<220>  
 <221> CDS  
 <222> (1457)..(4486)

<400> 1  
 ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggtca 60  
 gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaaggcagt  
 tcctgccccg gctcagggcc aagaacagat ggtccccaga tgcggtccag ccctcagcag 120  
 tttctagaga accatcagat gtttccaggg tgccccaagg acctgaaatg accctgtgcc 180  
 ttatggaaac taaccaatca gttcgcttct cgcttctgtt cgccgcgttc tgctccccga  
 gctcaataaa agagcccaca acccctcact cggggcgcca gtcctccgat tgactgagtc 240  
 gccccgggtac ccgtgtatcc aataaacctt cttgcagttg catccgactt gtggtctcgc 300  
 tgttccttgg gaggggtctcc tctgagtgtat tgactacccg tcagcgggggg tctttcattt 360  
 gggggctcgt ccgggatcgg gagaccctg cccagggacc accgaccac caccgggagg 420  
 540

0123456789

|   |      |
|---|------|
| caagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgatttta   | 600  |
| tgcgcctgct tcggtaact agctctgtat ctggcggacc cgtggggaa                | 660  |
| ctgacgagtt ctgaacacccc ggccgcaacc ctgggagacg tcccaggac tttggggcc    | 720  |
| gttttgggg cccgacactga ggaaggggagt cgatgtggaa tccgaccccg tcaggatatg  | 780  |
| tgggtctgggt aggagacgag aacctaaaac agttcccgcc tccgtctgaa ttttgcctt   | 840  |
| cggttggaa ccgaagccgc gcgtcttgc tgctgcagca tcggtctgtg ttgtctgt       | 900  |
| ccgactgtgt ttctgtatgt gtctgaaaat tagggccaga ctgttaccac tcccttaagt   | 960  |
| ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccagtc ggtagatgtc   | 1020 |
| aagaagagac gttgggttac cttctgctc gcagaatggc caacctttaa cgtcgatgg     | 1080 |
| ccgcgagacg gcacctttaa ccgagaccc tcacccagg ttaagatcaa ggtctttca      | 1140 |
| cctggccgc atggacacccc agaccaggc ccctacatcg tgacctggaa agccttggct    | 1200 |
| tttgcacccctt ctccctgggt caagccctt gtacacccta agcctccgc tcctttccct   | 1260 |
| ccatccgcgc cgtctctccc ctttgaacct cctcggtcgca cccgcctcg atcctccctt   | 1320 |
| tatccagccc tcactccttc tctaggcgcc ggccgctcta gcccattaat acgactca     | 1380 |
| atagggcgat tcgaatcagg ctttggcgcg ccggatcctt aattaagcgc aattgggagg   | 1440 |
| tggcggttagc ctgcgag atg ggc gtg att acg gat tca ctg gcc gtc gtg gcc | 1492 |
| Met Gly Val Ile Thr Asp Ser Leu Ala Val Val Ala                     |      |
| 1 5 10  |      |
| cgc acc gat cgc cct tcc caa cag tta cgc agc ctg aat ggc gaa tgg     | 1540 |
| Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp     |      |
| 15 20 25  |      |
| cgc ttt gcc tgg ttt ccg gca cca gaa gcg gtg ccg gaa agc tgg ctg     | 1588 |
| Arg Phe Ala Trp Phe Pro Ala Pro Glu Ala Val Pro Glu Ser Trp Leu     |      |
| 30 35 40  |      |
| gag tgc gat ctt cct gag gcc gat act gtc gtc gtc ccc tca aac tgg     | 1636 |
| Glu Cys Asp Leu Pro Glu Ala Asp Thr Val Val Val Pro Ser Asn Trp     |      |
| 45 50 55 60   |      |
| cag atg cac ggt tac gat gcg ccc atc tac acc aac gtg acc tat ccc     | 1684 |
| Gln Met His Gly Tyr Asp Ala Pro Ile Tyr Thr Asn Val Thr Tyr Pro     |      |
| 65 70 75  |      |
| att acg gtc aat ccg ccg ttt gtt ccc acg gag aat ccg acg ggt tgt     | 1732 |
| Ile Thr Val Asn Pro Pro Phe Val Pro Thr Glu Asn Pro Thr Gly Cys     |      |
| 80 85 90  |      |
| tac tcg ctc aca ttt aat gtt gat gaa agc tgg cta cag gaa ggc cag     | 1780 |
| Tyr Ser Leu Thr Phe Asn Val Asp Glu Ser Trp Leu Gln Glu Gly Gln     |      |
| 95 100 105  |      |

acg cga att att ttt gat ggc gtt aac tcg gcg ttt cat ctg tgg tgc 1828  
 Thr Arg Ile Ile Phe Asp Gly Val Asn Ser Ala Phe His Leu Trp Cys  
 110 115 120

aac ggg cgc tgg gtc ggt tac ggc cag gac agt cgt ttg ccg tct gaa 1876  
 Asn Gly Arg Trp Val Gly Tyr Gly Gln Asp Ser Arg Leu Pro Ser Glu  
 125 130 135 140

ttt gac ctg agc gca ttt tta cgc gcc gga gaa aac cgc ctc gcg gtg 1924  
 Phe Asp Leu Ser Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu Ala Val  
 145 150 155

atg gtg ctg cgc tgg agt gac ggc agt tat ctg gaa gat cag gat atg 1972  
 Met Val Leu Arg Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln Asp Met  
 160 165 170

tgg cgg atg agc ggc att ttc cgt gac gtc tcg ttg ctg cat aaa ccg 2020  
 Trp Arg Met Ser Gly Ile Phe Arg Asp Val Ser Leu Leu His Lys Pro  
 175 180 185

act aca caa atc agc gat ttc cat gtt gcc act cgc ttt aat gat gat 2068  
 Thr Thr Gln Ile Ser Asp Phe His Val Ala Thr Arg Phe Asn Asp Asp  
 190 195 200

ttc agc cgc gct gta ctg gag gct gaa gtt cag atg tgc ggc gag ttg 2116  
 Phe Ser Arg Ala Val Leu Glu Ala Glu Val Gln Met Cys Gly Glu Leu  
 205 210 215 220

cgt gac tac cta cgg gta aca gtt tct tta tgg cag ggt gaa acg cag 2164  
 Arg Asp Tyr Leu Arg Val Thr Val Ser Leu Trp Gln Gly Glu Thr Gln  
 225 230 235

gtc gcc agc ggc acc gcg cct ttc ggc ggt gaa att atc gat gag cgt 2212  
 Val Ala Ser Gly Thr Ala Pro Phe Gly Gly Glu Ile Ile Asp Glu Arg  
 240 245 250

ggt ggt tat gcc gat cgc gtc aca cta cgt ctg aac gtc gaa aac ccg 2260  
 Gly Gly Tyr Ala Asp Arg Val Thr Leu Arg Leu Asn Val Glu Asn Pro  
 255 260 265

aaa ctg tgg agc gcc gaa atc ccg aat ctc tat cgt gcg gtg gtt gaa 2308  
 Lys Leu Trp Ser Ala Glu Ile Pro Asn Leu Tyr Arg Ala Val Val Glu  
 270 275 280

ctg cac acc gcc gac ggc acg ctg att gaa gca gaa gcc tgc gat gtc 2356  
 Leu His Thr Ala Asp Gly Thr Leu Ile Glu Ala Glu Ala Cys Asp Val  
 285 290 295 300

ggt ttc cgc gag gtg cgg att gaa aat ggt ctg ctg ctg aac ggc 2404  
 Gly Phe Arg Glu Val Arg Ile Glu Asn Gly Leu Leu Leu Asn Gly  
 305 310 315

aag ccg ttg ctg att cga ggc gtt aac cgt cac gag cat cat cct ctg 2452  
 Lys Pro Leu Leu Ile Arg Gly Val Asn Arg His Glu His His Pro Leu  
 320 325 330

cat ggt cag gtc atg gat gag cag acg atg gtg cag gat atc ctg ctg 2500

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| His | Gly | Gln | Val | Met | Asp | Glu | Gln | Thr | Met | Val | Gln | Asp | Ile | Leu | Leu |      |
| 335 |     |     |     |     |     | 340 |     |     |     |     | 345 |     |     |     |     |      |
| atg | aag | cag | aac | aac | ttt | aac | gcc | gtg | cgc | tgt | tcg | cat | tat | ccg | aac | 2548 |
| Met | Lys | Gln | Asn | Asn |     | Phe | Asn | Ala | Val | Arg | Cys | Ser | His | Tyr | Pro | Asn  |
| 350 |     |     |     |     |     | 355 |     |     |     |     | 360 |     |     |     |     |      |
| cat | ccg | ctg | tgg | tac | acg | ctg | tgc | gac | cgc | tac | ggc | ctg | tat | gtg | gtg | 2596 |
| His | Pro | Leu | Trp | Tyr | Thr | Leu | Cys | Asp | Arg | Tyr | Gly | Leu | Tyr | Val | Val |      |
| 365 |     |     |     |     |     | 370 |     |     |     |     | 375 |     |     |     | 380 |      |
| gat | gaa | gcc | aat | att | gaa | acc | cac | ggc | atg | gtg | cca | atg | aat | cgt | ctg | 2644 |
| Asp | Glu | Ala | Asn | Ile | Glu | Thr | His | Gly | Met | Val | Pro | Met | Asn | Arg | Leu |      |
|     |     |     |     |     |     | 385 |     |     |     |     | 390 |     |     |     | 395 |      |
| acc | gat | gat | ccg | cgc | tgg | cta | ccg | gcf | atg | gtg | acc | cgc | gta | acg | cga | 2692 |
| Thr | Asp | Asp | Pro | Arg | Trp | Leu | Pro | Ala | Met | Ser | Glu | Arg | Val | Thr | Arg |      |
|     |     |     |     |     |     | 400 |     |     |     |     | 405 |     |     |     | 410 |      |
| atg | gtg | cag | cgc | gat | cgt | aat | cac | ccg | agt | gtg | atc | atc | tgg | tcg | ctg | 2740 |
| Met | Val | Gln | Arg | Asp | Arg | Asn | His | Pro | Ser | Val | Ile | Ile | Trp | Ser | Leu |      |
|     |     |     |     |     |     | 415 |     |     |     |     | 420 |     |     |     | 425 |      |
| ggg | aat | gaa | tca | ggc | cac | ggc | gct | aat | cac | gac | gcf | ctg | tat | cgc | tgg | 2788 |
| Gly | Asn | Glu | Ser | Gly | His | Gly | Ala | Asn | His | Asp | Ala | Leu | Tyr | Arg | Trp |      |
|     |     |     |     |     |     | 430 |     |     |     |     | 435 |     |     |     | 440 |      |
| atc | aaa | tct | gtc | gat | cct | tcc | cgc | ccg | gtg | cag | tat | gaa | ggc | ggc | gga | 2836 |
| Ile | Lys | Ser | Val | Asp | Pro | Ser | Arg | Pro | Val | Gln | Tyr | Glu | Gly | Gly |     |      |
|     |     |     |     |     |     | 445 |     |     |     |     | 450 |     |     |     | 460 |      |
| gcc | gac | acc | acg | gcc | acc | gat | att | att | tgc | ccg | atg | tac | gcf | cgc | gtg | 2884 |
| Ala | Asp | Thr | Ala | Thr | Asp | Ile | Ile | Cys | Pro | Met | Tyr | Ala | Arg | Val |     |      |
|     |     |     |     |     |     | 465 |     |     |     |     | 470 |     |     |     | 475 |      |
| gat | gaa | gac | cag | ccc | ttc | ccg | gct | gtg | ccg | aaa | tgg | tcc | atc | aaa | aaa | 2932 |
| Asp | Glu | Asp | Gln | Pro | Phe | Pro | Ala | Val | Pro | Lys | Trp | Ser | Ile | Lys | Lys |      |
|     |     |     |     |     |     | 480 |     |     |     |     | 485 |     |     |     | 490 |      |
| tgg | ctt | tcg | cta | cct | gga | gag | acg | cgc | ccg | ctg | atc | ctt | tgc | gaa | tac | 2980 |
| Trp | Leu | Ser | Leu | Pro | Gly | Glu | Thr | Arg | Pro | Leu | Ile | Leu | Cys | Glu | Tyr |      |
|     |     |     |     |     |     | 495 |     |     |     |     | 500 |     |     |     | 505 |      |
| gcc | cac | gcf | atg | ggt | aac | agt | ctt | ggc | ggt | ttc | gct | aaa | tac | tgg | cag | 3028 |
| Ala | His | Ala | Met | Gly | Asn | Ser | Leu | Gly | Gly | Phe | Ala | Lys | Tyr | Trp | Gln |      |
|     |     |     |     |     |     | 510 |     |     |     |     | 515 |     |     |     | 520 |      |
| gcf | ttt | cgt | cag | tat | ccc | cgt | tta | cag | ggc | ggc | ttc | gtc | tgg | gac | tgg | 3076 |
| Ala | Phe | Arg | Gln | Tyr | Pro | Arg | Leu | Gln | Gly | Gly | Phe | Val | Trp | Asp | Trp |      |
|     |     |     |     |     |     | 525 |     |     |     |     | 530 |     |     |     | 540 |      |
| gtg | gat | cag | tcg | ctg | att | aaa | tat | gat | gaa | aac | ggc | aac | ccg | tgg | tcg | 3124 |
| Val | Asp | Gln | Ser | Leu | Ile | Lys | Tyr | Asp | Glu | Asn | Gly | Asn | Pro | Trp | Ser |      |
|     |     |     |     |     |     | 545 |     |     |     |     | 550 |     |     |     | 555 |      |
| gct | tac | ggc | ggt | gat | ttt | ggc | gat | acg | ccg | aac | gat | cgc | cag | ttc | tgt | 3172 |
| Ala | Tyr | Gly | Gly | Asp | Phe | Gly | Asp | Thr | Pro | Asn | Asp | Arg | Gln | Phe | Cys |      |

560

565

570

|   |     |     |      |
|---|-----|-----|------|
| atg aac ggt ctg gtc ttt gcc gac cgc acg ccg cat cca gcg ctg acg<br>Met Asn Gly Leu Val Phe Ala Asp Arg Thr Pro His Pro Ala Leu Thr<br>575 | 580 | 585 | 3220 |
| gaa gca aaa cac cag cag ttt ttc cag ttc cgt tta tcc ggg caa<br>Glu Ala Lys His Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly Gln<br>590         | 595 | 600 | 3268 |
| acc atc gaa gtg acc agc gaa tac ctg ttc cgt cat agc gat aac gag<br>Thr Ile Glu Val Thr Ser Glu Tyr Leu Phe Arg His Ser Asp Asn Glu<br>605 | 610 | 615 | 3316 |
| ctc ctg cac tgg atg gtg gcg ctg gat ggt aag ccg ctg gca agc ggt<br>Leu Leu His Trp Met Val Ala Leu Asp Gly Lys Pro Leu Ala Ser Gly<br>625 | 630 | 635 | 3364 |
| gaa gtg cct ctg gat gtc gct cca caa ggt aaa cag ttg att gaa ctg<br>Glu Val Pro Leu Asp Val Ala Pro Gln Gly Lys Gln Leu Ile Glu Leu<br>640 | 645 | 650 | 3412 |
| cct gaa cta ccg cag ccg gag agc gcc ggg caa ctc tgg ctc aca gta<br>Pro Glu Leu Pro Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr Val<br>655 | 660 | 665 | 3460 |
| cgc gta gtg caa ccg aac gcg acc gca tgg tca gaa gcc ggg cac atc<br>Arg Val Val Gln Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly His Ile<br>670 | 675 | 680 | 3508 |
| agc gcc tgg cag cag tgg cgt ctg gcg gaa aac ctc agt gtg acg ctc<br>Ser Ala Trp Gln Gln Trp Arg Leu Ala Glu Asn Leu Ser Val Thr Leu<br>685 | 690 | 695 | 3556 |
| ccc gcc gcg tcc cac gcc atc ccg cat ctg acc acc agc gaa atg gat<br>Pro Ala Ala Ser His Ala Ile Pro His Leu Thr Thr Ser Glu Met Asp<br>705 | 710 | 715 | 3604 |
| ttt tgc atc gag ctg ggt aat aag cgt tgg caa ttt aac cgc cag tca<br>Phe Cys Ile Glu Leu Gly Asn Lys Arg Trp Gln Phe Asn Arg Gln Ser<br>720 | 725 | 730 | 3652 |
| ggc ttt ctt tca cag atg tgg att ggc gat aaa aaa caa ctg ctg acg<br>Gly Phe Leu Ser Gln Met Trp Ile Gly Asp Lys Gln Leu Leu Thr<br>735     | 740 | 745 | 3700 |
| ccg ctg cgc gat cag ttc acc cgt gca ccg ctg gat aac gac att ggc<br>Pro Leu Arg Asp Gln Phe Thr Arg Ala Pro Leu Asp Asn Asp Ile Gly<br>750 | 755 | 760 | 3748 |
| gta agt gaa gcg acc cgc att gac cct aac gcc tgg gtc gaa cgc tgg<br>Val Ser Glu Ala Thr Arg Ile Asp Pro Asn Ala Trp Val Glu Arg Trp<br>765 | 770 | 775 | 3796 |
| aag gcg gcg ggc cat tac cag gcc gaa gca gcg ttg ttg cag tgc acg<br>Lys Ala Ala Gly His Tyr Gln Ala Glu Ala Ala Leu Leu Gln Cys Thr<br>785 | 790 | 795 | 3844 |

gca gat aca ctt gct gat gcg gtg ctg att acg acc gct cac gcg tgg 3892  
 Ala Asp Thr Leu Ala Asp Ala Val Leu Ile Thr Thr Ala His Ala Trp  
 800 805 810

cag cat cag ggg aaa acc tta ttt atc agc cg<sup>g</sup> aaa acc tac cg<sup>g</sup> att 3940  
 Gln His Gln Gly Lys Thr Leu Phe Ile Ser Arg Lys Thr Tyr Arg Ile  
 815 820 825

gat ggt agt ggt caa atg gcg att acc gtt gat gtt gaa gtg gcg agc 3988  
 Asp Gly Ser Gly Gln Met Ala Ile Thr Val Asp Val Glu Val Ala Ser  
 830 835 840

gat aca ccg cat ccg gcg att ggc ctg aac tgc cag ctg gcg cag 4036  
 Asp Thr Pro His Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala Gln  
 845 850 855 860

gta gca gag cg<sup>g</sup> gta aac tgg ctc gga tta ggg cg<sup>g</sup> caa gaa aac tat 4084  
 Val Ala Glu Arg Val Asn Trp Leu Gly Leu Gly Pro Gln Glu Asn Tyr  
 865 870 875

ccc gac cgc ctt act gcc gcc tgg ttt gac cgc tgg gat ctg cca ttg 4132  
 Pro Asp Arg Leu Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu Pro Leu  
 880 885 890

tca gac atg tat acc ccg tac gtc ttc ccg agc gaa aac ggt ctg cgc 4180  
 Ser Asp Met Tyr Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly Leu Arg  
 895 900 905

tgc ggg acg cgc gaa ttg aat tat ggc cca cac cag tgg cgc ggc gac 4228  
 Cys Gly Thr Arg Glu Leu Asn Tyr Gly Pro His Gln Trp Arg Gly Asp  
 910 915 920

ttc cag ttc aac atc agc cgc tac agt caa cag caa ctg atg gaa acc 4276  
 Phe Gln Phe Asn Ile Ser Arg Tyr Ser Gln Gln Gln Leu Met Glu Thr  
 925 930 935 940

agc cat cgc cat ctg ctg cac gcg gaa gaa ggc aca tgg ctg aat atc 4324  
 Ser His Arg His Leu Leu His Ala Glu Glu Gly Thr Trp Leu Asn Ile  
 945 950 955

gac ggt ttc cat atg ggg att ggt ggc gac gac tcc tgg agc ccg tca 4372  
 Asp Gly Phe His Met Gly Ile Gly Gly Asp Asp Ser Trp Ser Pro Ser  
 960 965 970

gta tcg gcg gaa ttc cag ctg agc gcc ggt cgc tac cat tac cag ttg 4420  
 Val Ser Ala Glu Phe Gln Leu Ser Ala Gly Arg Tyr His Tyr Gln Leu  
 975 980 985

gtc tgg tgt caa aaa aga tct gac tat aaa gat gag gac ctc gac cat 4468  
 Val Trp Cys Gln Lys Arg Ser Asp Tyr Lys Asp Glu Asp Leu Asp His  
 990 995 1000

cat cat cat cac cg<sup>g</sup> taataatagg tagataagtg actgattaga 4516  
 His His His His Arg  
 1005 1010

tgcattgatc cctcgaccaa ttccggttat ttccaccat attgccgtct ttggcaatg 4576

tgagggcccg gaaacctggc cctgtttct tgacgagcat tccttaggggt cttccccctc 4636  
 tcgccaagg aatgcaaggt ctgttgaatg tcgtgaagga agcagttcct ctggaagctt 4696  
 cttgaagaca aacaacgtct gtagcgaccc tttgcaggca gcggaacccc ccacctggcg 4756  
 acaggtgcct ctgcggccaa aagccacgtg tataagatac acctgcaaag gcggcacaac 4816  
 cccagtgcga cggtgtgagt tggatagttg tggaaagagt caaatggctc tcctcaagcg 4876  
 tattcaacaa ggggctgaag gatgcccaga aggtacccca ttgtatgggat tctgatctgg 4936  
 ggccctcggtg cacatgctt acatgtgttt agtcgagggtt aaaaaacgtc taggcccccc 4996  
 gaaccacggg gacgtggttt tcctttgaaa aacacgatga taataccatg attgaacaag 5056  
 atggattgca cgccagggttcc cgccgcgtt gggtggagag gctattcggc tatgactggg 5116  
 cacaacagac aatcggtcgc tctgatgccc ccgtgttccg gctgtcagcg cagggcgcc 5176  
 cggttctttt tgtcaagacc gacctgtccg gtgcctgaa tgaactgcag gacgaggcag 5236  
 cgccgctatc gtggctggcc acgacggggcg ttccctgcgc agctgtgctc gacgttgtca 5296  
 ctgaagcggg aaggactgg ctgctattgg gcgaagtgcc ggggcaggat ctcctgtcat 5356  
 ctcacccgtc tcctgcccgg aaagtatcca tcatggctga tgcaatgcgg cggctgcata 5416  
 cgcttgcattcc ggctacccgtc ccattcgacc accaagcgaa acatcgcatc gagcgagcac 5476  
 gtactcgat ggaagccggc cttgtcgatc aggtatctt ggacgaagag catcaggggc 5536  
 tcgcgccagc cgaactgttc gccaggctca aggccgcattt gcccgcacggc gaggatctcg 5596  
 tcgtgaccca tggcgatgcc tgcttgcga atatcatggt ggaaaatggc cgctttctg 5656  
 gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggcta 5716  
 cccgtgatat tgctgaagag cttggcgccg aatggctga ccgcttcctc gtgtttacg 5776  
 gtatcgccgc tcccgattcg cagcgcatcg cttctatcg cttcttgac gagttttct 5836  
 gagcggact ctggggttcg catcgataaa ataaaagatt ttattnatgc tccagaaaaa 5896  
 gggggaaatg aaagacccca cctgttaggtt tggcaagcta gcttaagtaa cgccattttg 5956  
 caaggcatgg aaaaatacat aactgagaat agagaagttc agatcaaggt caggaacaga 6016  
 tggaaacagct gaatatggc caaacaggat atctgtggta agcagttcct gccccggctc 6076  
 agggccaaga acagatggaa cagctgaata tggccaaac aggatatctg tggtaagcag 6136  
 ttccctgcccc ggctcagggc caagaacaga tggccccag atgcggtcca gccctcagca 6196  
 gtttctagag aaccatcaga tgtttccagg gtgccccaaag gacctgaaat gaccctgtgc 6256  
 cttatggaa ctaaccaatc agttcgcttc tcgcttctgt tcgcgcgctt ctgctcccc 6316

|   |      |
|---|------|
| agctcaataa aagagccac aaccctcac tcggggcgcc agtcctccga ttgactgagt   | 6376 |
| cgcgggtta cccgtgtatc caataaaccc tcttcagtt gcatccgact tgtggtctcg   | 6436 |
| ctgttccttgcgggtctc ctctgagtga ttgactaccc gtcagcgggg gtcttcatt     | 6496 |
| catgcagcat gtatcaaaat taatttggtt tttttctta agtatttaca ttaaatggcc  | 6556 |
| atagttgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttggcgctct | 6616 |
| tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca | 6676 |
| gctcactcaa aggccgtaat acgg  | 6700 |

<210> 2  
<211> 1010  
<212> PRT  
<213> Unknown

<220>  
<223> pICAST ALC.

<400> 2

|   |    |    |    |
|---|----|----|----|
| Met Gly Val Ile Thr Asp Ser Leu Ala Val Val Ala Arg Thr Asp Arg |    |    |    |
| 1   | 5  | 10 | 15 |
| 10  | 15 |    |    |

|   |    |    |  |
|---|----|----|--|
| Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp Arg Phe Ala Trp |    |    |  |
| 20  | 25 | 30 |  |
| 30  |    |    |  |

|   |    |    |  |
|---|----|----|--|
| Phe Pro Ala Pro Glu Ala Val Pro Glu Ser Trp Leu Glu Cys Asp Leu |    |    |  |
| 35  | 40 | 45 |  |
| 45  |    |    |  |

|   |    |    |  |
|---|----|----|--|
| Pro Glu Ala Asp Thr Val Val Pro Ser Asn Trp Gln Met His Gly |    |    |  |
| 50  | 55 | 60 |  |
| 60  |    |    |  |

|   |    |    |    |
|---|----|----|----|
| Tyr Asp Ala Pro Ile Tyr Thr Asn Val Thr Tyr Pro Ile Thr Val Asn |    |    |    |
| 65  | 70 | 75 | 80 |
| 75  | 80 |    |    |

|   |    |    |  |
|---|----|----|--|
| Pro Pro Phe Val Pro Thr Glu Asn Pro Thr Gly Cys Tyr Ser Leu Thr |    |    |  |
| 85  | 90 | 95 |  |
| 95  |    |    |  |

|   |     |     |  |
|---|-----|-----|--|
| Phe Asn Val Asp Glu Ser Trp Leu Gln Glu Gly Gln Thr Arg Ile Ile |     |     |  |
| 100   | 105 | 110 |  |
| 110   |     |     |  |

|   |     |     |  |
|---|-----|-----|--|
| Phe Asp Gly Val Asn Ser Ala Phe His Leu Trp Cys Asn Gly Arg Trp |     |     |  |
| 115   | 120 | 125 |  |
| 125   |     |     |  |

Val Gly Tyr Gly Gln Asp Ser Arg Leu Pro Ser Glu Phe Asp Leu Ser  
 130 135 140

Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu Ala Val Met Val Leu Arg  
 145 150 155 160

Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln Asp Met Trp Arg Met Ser  
 165 170 175

Gly Ile Phe Arg Asp Val Ser Leu Leu His Lys Pro Thr Thr Gln Ile  
 180 185 190

Ser Asp Phe His Val Ala Thr Arg Phe Asn Asp Asp Phe Ser Arg Ala  
 195 200 205

Val Leu Glu Ala Glu Val Gln Met Cys Gly Glu Leu Arg Asp Tyr Leu  
 210 215 220

Arg Val Thr Val Ser Leu Trp Gln Gly Glu Thr Gln Val Ala Ser Gly  
 225 230 235 240

Thr Ala Pro Phe Gly Gly Glu Ile Ile Asp Glu Arg Gly Gly Tyr Ala  
 245 250 255

Asp Arg Val Thr Leu Arg Leu Asn Val Glu Asn Pro Lys Leu Trp Ser  
 260 265 270

Ala Glu Ile Pro Asn Leu Tyr Arg Ala Val Val Glu Leu His Thr Ala  
 275 280 285

Asp Gly Thr Leu Ile Glu Ala Glu Ala Cys Asp Val Gly Phe Arg Glu  
 290 295 300

Val Arg Ile Glu Asn Gly Leu Leu Leu Asn Gly Lys Pro Leu Leu  
 305 310 315 320

Ile Arg Gly Val Asn Arg His Glu His His Pro Leu His Gly Gln Val  
 325 330 335

Met Asp Glu Gln Thr Met Val Gln Asp Ile Leu Leu Met Lys Gln Asn  
 340 345 350

Asn Phe Asn Ala Val Arg Cys Ser His Tyr Pro Asn His Pro Leu Trp

355

360

365

Tyr Thr Leu Cys Asp Arg Tyr Gly Leu Tyr Val Val Asp Glu Ala Asn  
 370 375 380

Ile Glu Thr His Gly Met Val Pro Met Asn Arg Leu Thr Asp Asp Pro  
 385 390 395 400

Arg Trp Leu Pro Ala Met Ser Glu Arg Val Thr Arg Met Val Gln Arg  
 405 410 415

Asp Arg Asn His Pro Ser Val Ile Ile Trp Ser Leu Gly Asn Glu Ser  
 420 425 430

Gly His Gly Ala Asn His Asp Ala Leu Tyr Arg Trp Ile Lys Ser Val  
 435 440 445

Asp Pro Ser Arg Pro Val Gln Tyr Glu Gly Gly Ala Asp Thr Thr  
 450 455 460

Ala Thr Asp Ile Ile Cys Pro Met Tyr Ala Arg Val Asp Glu Asp Gln  
 465 470 475 480

Pro Phe Pro Ala Val Pro Lys Trp Ser Ile Lys Lys Trp Leu Ser Leu  
 485 490 495

Pro Gly Glu Thr Arg Pro Leu Ile Leu Cys Glu Tyr Ala His Ala Met  
 500 505 510

Gly Asn Ser Leu Gly Gly Phe Ala Lys Tyr Trp Gln Ala Phe Arg Gln  
 515 520 525

Tyr Pro Arg Leu Gln Gly Gly Phe Val Trp Asp Trp Val Asp Gln Ser  
 530 535 540

Leu Ile Lys Tyr Asp Glu Asn Gly Asn Pro Trp Ser Ala Tyr Gly Gly  
 545 550 555 560

Asp Phe Gly Asp Thr Pro Asn Asp Arg Gln Phe Cys Met Asn Gly Leu  
 565 570 575

Val Phe Ala Asp Arg Thr Pro His Pro Ala Leu Thr Glu Ala Lys His  
 580 585 590

Gln Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly Gln Thr Ile Glu Val  
 595 600 605

Thr Ser Glu Tyr Leu Phe Arg His Ser Asp Asn Glu Leu Leu His Trp  
 610 615 620

Met Val Ala Leu Asp Gly Lys Pro Leu Ala Ser Gly Glu Val Pro Leu  
 625 630 635 640

Asp Val Ala Pro Gln Gly Lys Gln Leu Ile Glu Leu Pro Glu Leu Pro  
 645 650 655

Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr Val Arg Val Val Gln  
 660 665 670

Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly His Ile Ser Ala Trp Gln  
 675 680 685

Gln Trp Arg Leu Ala Glu Asn Leu Ser Val Thr Leu Pro Ala Ala Ser  
 690 695 700

His Ala Ile Pro His Leu Thr Thr Ser Glu Met Asp Phe Cys Ile Glu  
 705 710 715 720

Leu Gly Asn Lys Arg Trp Gln Phe Asn Arg Gln Ser Gly Phe Leu Ser  
 725 730 735

Gln Met Trp Ile Gly Asp Lys Lys Gln Leu Leu Thr Pro Leu Arg Asp  
 740 745 750

Gln Phe Thr Arg Ala Pro Leu Asp Asn Asp Ile Gly Val Ser Glu Ala  
 755 760 765

Thr Arg Ile Asp Pro Asn Ala Trp Val Glu Arg Trp Lys Ala Ala Gly  
 770 775 780

His Tyr Gln Ala Glu Ala Ala Leu Leu Gln Cys Thr Ala Asp Thr Leu  
 785 790 795 800

Ala Asp Ala Val Leu Ile Thr Thr Ala His Ala Trp Gln His Gln Gly  
 805 810 815

Lys Thr Leu Phe Ile Ser Arg Lys Thr Tyr Arg Ile Asp Gly Ser Gly  
 820 825 830

Gln Met Ala Ile Thr Val Asp Val Glu Val Ala Ser Asp Thr Pro His  
 835 840 845

Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala Gln Val Ala Glu Arg  
 850 855 860

Val Asn Trp Leu Gly Leu Gly Pro Gln Glu Asn Tyr Pro Asp Arg Leu  
 865 870 875 880

Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu Pro Leu Ser Asp Met Tyr  
 885 890 895

Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly Leu Arg Cys Gly Thr Arg  
 900 905 910

Glu Leu Asn Tyr Gly Pro His Gln Trp Arg Gly Asp Phe Gln Phe Asn  
 915 920 925

Ile Ser Arg Tyr Ser Gln Gln Leu Met Glu Thr Ser His Arg His  
 930 935 940

Leu Leu His Ala Glu Glu Gly Thr Trp Leu Asn Ile Asp Gly Phe His  
 945 950 955 960

Met Gly Ile Gly Gly Asp Asp Ser Trp Ser Pro Ser Val Ser Ala Glu  
 965 970 975

Phe Gln Leu Ser Ala Gly Arg Tyr His Tyr Gln Leu Val Trp Cys Gln  
 980 985 990

Lys Arg Ser Asp Tyr Lys Asp Glu Asp Leu Asp His His His His His  
 995 1000 1005

His Arg  
 1010

<210> 3  
 <211> 6700  
 <212> DNA  
 <213> Unknown

<220>  
<223> pICAST ALC.

<400> 3  
 gacgtcggac ttatacccg tttgtcctat agacaccatt cgtcaaggac ggggccgagt 60  
 cccggttctt gtctacccgt tcgacttata cccgggttgc cctatagaca ccattcgta. 120  
 aggacggggc cgagtcccg ttttgtcta ccaggggtct acgccaggc gggagtcgtc 180  
 aaagatctct tggtagtcta caaaggccc acggggttcc tggactttac tggacacgg 240  
 aataaaacttg attggtagt caagcgaaga gcgaagacaa gcgcgcgaag acgaggggt 300  
 cgagttattt tctcgggtgt tggggagtga gccccgcgtt caggaggcta actgactcag 360  
 cggcccatg ggcacatagg ttatggga gaacgtcaac gtggctgaa caccagacgc 420  
 acaaggaacc ctcccagagg agactcacta actgatggc agtcgcccc agaaagtaaa 480  
 ccccgagca ggccctagcc ctctgggac gggccctgg tggctgggtg gtggccctcc 540  
 gttcgaccgg tcgttgaata gacacagaca ggctaacaga tcacagatac tgactaaaat 600  
 acgcggacgc agccatgatc aatcgattga tcgagacata gaccgcctgg gcaccacatt 660  
 gactgctcaa gacttgggg ccggcggtgg gaccctctgc agggtccctg aaaccccccgg 720  
 caaaaacacc gggctggact cttccctca gctacacctt aggctgggc agtccctatac 780  
 accaagacca tcctctgctc ttggatttttgc tcaagggcgg aggcaagactt aaaaacgaaa 840  
 gccaaacacc ttggatggggcc cgcagaacag acgacgtcgt agcaagacac aacagagaca 900  
 gactgacaca aagacataaa cagacttttgc atcccggtct gacaatggtg aggaaattca 960  
 aacttggaaatc cattgacccctt tctacagctc gccgagcgtt gttgggttagt ccatctacag 1020  
 ttcttctctg caacccaaatg gaagacgaga cgtcttaccg gttggaaattt gcagcctacc 1080  
 ggcgtctgc cgtggaaattt ggctctggag tagtgggtcc aattctagtt ccagaaaagt 1140  
 ggaccggggcg tacctgtggg tctggtccag gggatgttagc actggaccct tcggaaaccga 1200  
 aaactggggg gagggaccca gttcgggaaa catgtggat tcggaggcgg aggagaagga 1260  
 ggtaggcggg gcagagaggg ggaacttggg ggagcaagct gggcggagc taggaggaa 1320  
 ataggcggg agtgaggaag agatcccgcc cggcggat cggtaattt tgctgagtga 1380  
 tatcccgcta agcttagtcc ggaaccgcgc ggcctaggaa ttaattcgct ttaaccctcc 1440  
 accggccatcg gagctctacc cgcactaatg cctaaatgc cggcggat cggcgtggct 1500  
 agcgggaagg gttgtcaatg cgtcggactt accgcttacc gcgaaacggc ccaaaggccg 1560  
 tggcttcgc cacggcctt cgaccgaccc cacgctagaa ggactccggc tatgacagca 1620

|  |      |
|--|------|
| gcagggggagt ttgaccgtct acgtgccaat gctacgcggg tagatgtggt tgcactggat   | 1680 |
| aggtaatgc cagttaggcg gcaaacaagg gtgcctctta ggctgcccaa caatgagcga     | 1740 |
| gtgtaaatta caactacttt cgaccgatgt cttccggc tgcgttaat aaaaactacc       | 1800 |
| gcaattgagc cgcaaagtag acaccacgtt gcccgcgacc cagccatgc cggtcctgtc     | 1860 |
| agcaaacggc agacttaaac tggactcgcg taaaaatgcg cggcctctt tggcggagcg     | 1920 |
| ccactaccac gacgcgaccc cactgccgtc aatagacctt ctagtcctat acaccgccta    | 1980 |
| ctcgccgtaa aaggcactgc agagcaacga cgtatggc tgcgtgttt agtcgtaaa        | 2040 |
| ggtacaacgg tgagcgaat tactactaaa gtcggcgca catgacctcc gacttcaagt      | 2100 |
| ctacacgccc ctcaacgcac tgcgtgtgc ccattgtcaa agaaataccg tcccacttt      | 2160 |
| cgtccagcgg tcgcccgtgc gcggaaagcc gccacttta tagtactcg caccaccaat      | 2220 |
| acggctagcg cagtgtgatg cagacttgca gctttggc tttgacacct cgccgctta       | 2280 |
| ggccttagag atagcacgccc accaacttga cgtgtggcgg ctgcgtgcg actaacttcg    | 2340 |
| tcttcggacg ctacagccaa aggcgctcca cgcctaactt ttaccagacg acgacgactt    | 2400 |
| gccgttcggc aacgactaag ctccgcaatt ggcagtgcgc gtagtaggag acgtaccagt    | 2460 |
| ccagttaccta ctcgtctgct accacgtcct ataggacgac tacttcgtct tggtaatt     | 2520 |
| gcggcacgca acaagcgtaa taggctggt aggcgacacc atgtgcgaca cgctggcgat     | 2580 |
| gcgggacata caccacccat ttcggttata actttgggtg ccgtaccacg gttacttagc    | 2640 |
| agactggcta ctggcgca ccgatggccg ctactcgctt ggcattgcg cttaccacgt       | 2700 |
| cgcgttagca ttatgtggct cacactagta gaccagcgac cccttactta gtccgggtcc    | 2760 |
| gcgatttagtgc tgccgcgaca tagcgaccta gtttagacag ctaggaaggg cgggcccacgt | 2820 |
| catacttccg ccgcctcggc tgggtgccc gtggctataa taaacgggct acatgcgcgc     | 2880 |
| gcacctactt ctggtcggga agggccgaca cggcttacc aggtagttt ttaccgaaag      | 2940 |
| cgtggaccc ctctgcgcgg ggcacttagga aacgcttatg cgggtgcgc acccattgtc     | 3000 |
| agaaccgcac aagcgattta tgaccgtccg caaagcagtc ataggggcaaa atgtcccgc    | 3060 |
| gaagcagacc ctgacccacc tagtcagcga ctaatttata ctactttgc cttttggcac     | 3120 |
| cagccgaatg ccgcactaa aaccgctatg cggcttgcta gcggtcaaga catabtgc       | 3180 |
| agaccagaaa cggctggcgt gcccgttagg tcgcactgc cttcggtttg tggcgtcgt      | 3240 |
| caaaaaggc aaggcaaata ggccgtttt gtagcttac tggcgtctt tggacaaggc        | 3300 |
| atatcgcta ttgctcgagg acgtgaccta ccaccgcgac ctaccattcg gcgaccgttc     | 3360 |

|   |      |
|---|------|
| gccacttcac ggagacctac agcgagggtgt tccatttgc aactaacttg acggacttga   | 3420 |
| tggcgctggc ctctcgccgc ccgttgagac cgagtgtcat gcgcatcacg ttggcttgcg   | 3480 |
| ctggcgtaacc agtcttcggc ccgtgtagtc gcggaccgtc gtcaccgcag accgccttt   | 3540 |
| ggagtcacac tgcgaggggc ggcgcagggt gcggtagggc gtagactggt ggtcgcttta   | 3600 |
| cctaaaaacg tagctcgacc cattattcgc aaccgttaaa ttggcggtca gtccgaaaga   | 3660 |
| aagtgtctac acctaaccgc tatttttgc tgacgactgc ggcgacgcgc tagtcaagtgc   | 3720 |
| ggcacgtggc gacctattgc tgtaaccgc ttcacttcgc tggcgtaac tgggattgcg     | 3780 |
| gaccaggctt gcgaccttcc gccgccccgt aatggtccgg cttcgtcgca acaacgtcac   | 3840 |
| gtgccgtcta tgtgaacgac tacgcccacga ctaatgctgg cgagtgcgc ccgtcgtagt   | 3900 |
| ccccctttgg aataaaatagt cggccttttgc gatggcctaa ctaccatcac cagtttaccg | 3960 |
| ctaattggcaa ctacaacttc accgctcgct atgtggcgta ggccgcgcct aaccggactt  | 4020 |
| gacggtcgac cgctccatc gtctcgccca tttgaccgag cctaattcccg gcgttctttt   | 4080 |
| gatagggctg gcggaatgac ggcggacaaa actggcgacc ctagacggta acagtctgt    | 4140 |
| catatggggc atgcagaagg gctcgctttt gccagacgcg acgcccgtcg cgcttaactt   | 4200 |
| aataccgggt gtggtcacccg cggcgtgaa ggtcaagtttgc tagtcggcgttgtcagttgt  | 4260 |
| cgttgactac ctttggtcgg tagcggtaga cgacgtgcgc cttcttcgt gtaccgactt    | 4320 |
| atagctggca aaggtataacc cctaaccacc gctgctgagg acctcgccgc gtcatacgcc  | 4380 |
| ccttaaggtc gactcgccgc cagcgatggt aatggtcaac cagaccacag tttttcttag   | 4440 |
| actgatattt ctactcctgg agctggtagt agtagtagta gtggccatata ttatccatct  | 4500 |
| attcactgac taatctacgt aactagggag ctggtaagg ccaataaaag gtggtataac    | 4560 |
| ggcagaaaac cgttacactc cggggcccttt ggaccgggac agaagaactg ctcgtaaagg  | 4620 |
| tccccagaaa ggggagagcg gttccattac gttccagaca acttacagca cttccttcgt   | 4680 |
| caaggagacc ttcaagaac ttctgtttgt tgcaagacatc gctggaaac gtccgtcgcc    | 4740 |
| ttgggggggt gaccgctgtc cacggagacg ccggtttcg gtgcacatat tctatgtgga    | 4800 |
| cgttccgccc gtgttgggt cacggtgcaa cactcaaccc atcaacaccc ttctcgttt     | 4860 |
| accgagagga gttcgataa gttgttcccc gacttcctac gggtcttcca tggggtaaca    | 4920 |
| tacccttagac tagaccccg agccacgtgt acgaaatgtt cacaatcag ctccaaattt    | 4980 |
| ttgcagatcc ggggggcttg gtgcccgtc accaaaagga aacttttgc gctactatttt    | 5040 |
| ttgtactaac ttgttctacc taacgtgcgt ccaagaggcc ggcgaaccca cctctccgat   | 5100 |

aagccgatac tgaccctgtgt tgtctgttag ccgacgagac tacggcggca caaggccgac 5160  
 agtcgcgtcc ccgcgggcca agaaaaacag ttctggctgg acaggccacg ggacttactt 5220  
 gacgtcctgc tccgtcgccgc cgatagcacc gaccggtgct gcccgcagg aacgcgtcga 5280  
 cacgagctgc aacagtgact tcgccccttcc ctgaccgacg ataaccgcct tcacggcccc 5340  
 gtcctagagg acagtagagt ggaacgagga cggcttttc ataggttagta ccgactacgt 5400  
 tacgcccggc acgtatgcga actaggccga tggacgggta agctggtggt tcgctttgta 5460  
 gcgtagctcg ctcgtgcattg agcctacctt cggccagaac agctagtccct actagacctg 5520  
 cttctcgtag tccccgagcg cggtcggtt gacaagcggt ccgagttccg cgcgtacggg 5580  
 ctggcgctcc tagagcagca ctgggttaccg ctacggacga acggcttata gtaccacctt 5640  
 ttaccggcga aaagacctaa gtagctgaca cccggccgacc cacaccgcct ggcgatagtc 5700  
 ctgtatcgca accgatgggc actataacga cttctcgaaac cgccgcttac ccgactggcg 5760  
 aaggagcacg aaatgccata gcggcgaggg ctaagcgctcg ctagcgaa gatagcgaa 5820  
 gaactgctca agaagactcg ccctgagacc ccaagcgtag ctatttatt ttctaaaata 5880  
 aatcagaggt cttttcccc cttactttc tggggggac atccaaaccg ttcgatcgaa 5940  
 ttcattgcgg taaaacgttc cgtaccttt tatgtattga ctcttatctc ttcaagtcta 6000  
 gttccagtcc ttgtctacct tgcgactta tacccggttt gtcctataga caccattcgt 6060  
 caaggacggg gccgagtcgg ggttcttgtc tacctgtcg acttataccc ggttgcct 6120  
 atagacacca ttcgtaagg acggggccga gtcccggttc ttgtctacca ggggtctacg 6180  
 ccaggtcgaa agtcgtcaaa gatctcttg tagtctacaa aggtcccacg ggggtcctgg 6240  
 actttactgg gacacggaat aaacttgatt ggttagtcaa gcgaagagcg aagacaagcg 6300  
 cgcgaaagacg aggggctcga gttatcccctt cgggtgttgg ggagtgagcc ccgcggtcag 6360  
 gaggtctact gactcagcgg gcccattggc acataggta ttggggagaa cgtcaacgt 6420  
 ggctgaacac cagagcgaca aggaaccctc ccagaggaga ctcactaact gatgggcagt 6480  
 cgcggccaga aagtaagtac gtcgtacata gtttaatta aaccaaaaaa aagaattcat 6540  
 aaatgttaatt taccggtatac aacgttaatta cttagccgt tgcgccccc tctccgccaa 6600  
 acgcataacc gcgagaaggc gaaggagcga gtgactgagc gacgcgagcc agcaagccga 6660  
 cgcgcgtcgc catagtcgag tgagttccg ccattatgcc 6700

<210> 4  
 <211> 8518

<212> DNA  
 <213> Unknown

<220>  
 <223> pICAST ALN.

|   |      |
|---|------|
| <400> 4   |      |
| ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggtca          | 60   |
| gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaaggcagt        | 120  |
| tcctgccccg gctcaggggcc aagaacagat ggtccccaga tgcggtccag ccctcagcag        | 180  |
| tttctagaga accatcagat gtttccaggg tgccccaagg acctgaaatg accctgtgcc         | 240  |
| ttatattgaac taaccaatca gttcgcttct cgcttctgtt cgcgcgcttc tgctccccga        | 300  |
| gctcaataaa agagcccaca acccgtcact cggggcgcca gtcctccgat tgactgagtc         | 360  |
| gcccgggtac ccgtgtatcc aataaaccct cttgcagttg catccgactt gtggtctcgc         | 420  |
| tgttccttgg gagggtctcc tctgagtgat tgactacccg tcagcggggg tctttcattt         | 480  |
| gggggctcgt ccgggatcgg gagacccctg cccagggacc accgaccac caccggagg           | 540  |
| caagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgatttt          | 600  |
| tgcgccctgctc tcggtaact agctctgtat ctggcgacc cgtggtgaa                     | 660  |
| ctgacgagtt ctgaacaccc ggccgcaacc ctgggagacg tcccaggac tttggggcc           | 720  |
| gtttttgtgg cccgacctga ggaaggaggat cgatgtggaa tccgaccccg tcaggatatg        | 780  |
| tggttcttgtt aggagacgag aacctaaaac agttcccgcc tccgtctgaa tttttgcttt        | 840  |
| cggtttggaa ccgaagccgc gcgtcttgc tgctgcagca tcgttctgtt ttgtctctgt          | 900  |
| ctgactgtgt ttctgtat ttgtctgtt gtctgaaaat tagggccaga ctgttaccac tcccttaagt | 960  |
| ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccagtc ggtagatgtc         | 1020 |
| aagaagagac gttgggttac ctctgtct gcagaatggc caacctttaa cgtcggatgg           | 1080 |
| ccgcgagacg gcacctttaa ccgagacctc atcaccagg ttaagatcaa ggtctttca           | 1140 |
| cctggcccgcc atggacaccc agaccaggc ccctacatcg tgacctggaa agccttggct         | 1200 |
| tttgcacccccc ctccctgggt caagcccttt gtacacccta agcctccgccc tcctttcct       | 1260 |
| ccatcccgccc cgtctctccc ccttgaacct cctcggtcga ccccgccctcg atcctccctt       | 1320 |
| tatccagcccc tcactccttc tctaggcgcc ggccgctcta gcccattaat acgactcact        | 1380 |
| atagggcgat tcgaacacca tgcaccatca tcatacatcac gtcgactata aagatgagga        | 1440 |
| cctcgagatg ggcgtgatta cggattcact ggccgtcggt gcccgcacccg atcgcccttc        | 1500 |

DRAFT

|             |            |             |            |            |             |      |
|-------------|------------|-------------|------------|------------|-------------|------|
| ccaacagtta  | cgcagcctga | atggcgaatg  | gcgccttgcc | tggttccgg  | caccagaagc  | 1560 |
| ggtgcggaa   | agctggctgg | agtgcgatct  | tcctgaggcc | gatactgtcg | tcgtcccctc  | 1620 |
| aaactggcag  | atgcacggtt | acgatgcgcc  | catctacacc | aacgtgacct | atcccattac  | 1680 |
| ggtcaatccg  | cggtttgttc | ccacggagaa  | tccgacgggt | tgttactcgc | tcacattaa   | 1740 |
| tgttcatgaa  | agctggctac | aggaaggcca  | gacgcgaatt | atttttgatg | gcgttaactc  | 1800 |
| ggcggttcat  | ctgtggtgca | acgggcgctg  | ggtcggttac | ggccaggaca | gtcggttgcc  | 1860 |
| gtctgaattt  | gacctgagcg | catttttacg  | cgccggagaa | aaccgcctcg | cggtgatgg   | 1920 |
| gctgggctgg  | agtgacggca | gttatctgga  | agatcaggat | atgtggcgga | tgagcggcat  | 1980 |
| tttccgtgac  | gtctcggtgc | tgcataaaacc | gactacacaa | atcagcgatt | tccatgttgc  | 2040 |
| cactcgcttt  | aatgatgatt | rcagccgcgc  | tgtactggag | gctgaagttc | agatgtgcgg  | 2100 |
| cgagttgcgt  | gactacctac | ggtaacagt   | ttctttatgg | cagggtgaaa | cgcaggtcgc  | 2160 |
| cagcggcacc  | gcgccttctg | gcggtaaaat  | tatcgatgag | cgtggtggtt | atgcgcgtcg  | 2220 |
| cgtcacacta  | cgtctgaacg | tcgaaaaccc  | gaaactgtgg | agcgccgaaa | tcccgaatct  | 2280 |
| ctatcgtgcg  | gtgggtgaac | tgcacacccgc | cgacggcacg | ctgattgaag | cagaaggcctg | 2340 |
| cgtatgtcggt | ttcccgagg  | tgcggattga  | aatggtctg  | ctgctgctga | acggcaagcc  | 2400 |
| gttgctgatt  | cgaggcgta  | accgtcacga  | gcatcatcct | ctgcatggtc | aggtcatgga  | 2460 |
| tgagcagacg  | atgggtcagg | atatcctgct  | gatgaagcag | aacaacttta | acgcgcgtcg  | 2520 |
| ctgttcgcac  | tatccgaacc | atccgcgttg  | gtacacgctg | tgcgaccgct | acggcctgta  | 2580 |
| tgtggtgat   | gaagccaata | ttgaaaccca  | cggcatggtg | ccaatgaatc | gtctgaccga  | 2640 |
| tgtatccgcgc | tggctaccgg | cgatgagcga  | acgcgtacg  | cgaatggtgc | agcgcgatcg  | 2700 |
| taatcaccgg  | agtgtgatca | tctggtcgct  | gggaaatgaa | tcaggccacg | gcgcataatca | 2760 |
| cgacgcgcgt  | tatcgctgga | tcaaatactgt | cgatccttcc | cgcgggtgc  | agtatgaagg  | 2820 |
| cggcggagcc  | gacaccacgg | ccaccgatat  | tatttgcgg  | atgtacgcgc | gcgtggatga  | 2880 |
| agaccagccc  | ttcccggctg | tgccgaaatg  | gtccatcaaa | aaatggcttt | cgctacctgg  | 2940 |
| agagacgcgc  | ccgctgatcc | tttgcgata   | cgcccacgcg | atggtaaca  | gtcttggcg   | 3000 |
| tttcgctaaa  | tactggcagg | cgtttcgta   | gtatccccgt | ttacaggcg  | gcttcgtctg  | 3060 |
| ggactgggtg  | gatcagtcgc | tgattaaata  | tgtaaaaac  | ggcaaccgt  | ggtcggctta  | 3120 |
| cggcggtat   | tttggcgata | cgcgaacga   | tcgcagttc  | tgtatgaacg | gtctggctt   | 3180 |
| tgccgaccgc  | acgcccgcac | cagcgctgac  | ggaagcaaaa | caccagcagc | agtttttcca  | 3240 |

PCT/US2013/048450

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gttccgttta  | tccgggcaaa  | ccatcgaa    | gaccagcgaa  | tacctgttcc  | gtcatagcgaa | 3300 |
| taacgagctc  | ctgcactgga  | tggtggcgct  | ggatggtaag  | ccgctggcaa  | gcggtaagt   | 3360 |
| gcctctggat  | gtcgctccac  | aaggtaaaca  | gttgattgaa  | ctgcctgaac  | taccgcagcc  | 3420 |
| ggagagcgcc  | gggcaactct  | ggctcacagt  | acgcgttagt  | caaccgaacg  | cgaccgcatt  | 3480 |
| gtcagaagcc  | gggcacatca  | gcccctggca  | gcagtgccgt  | ctggcggaaa  | acctcagtgt  | 3540 |
| gacgctcccc  | gccgcgtccc  | acgccatccc  | gcatctgacc  | accagcgaaa  | tggatttttg  | 3600 |
| catcgagctg  | ggtaataaagc | gttggcaatt  | taaccgcag   | tcaggcttcc  | tttcacagat  | 3660 |
| gtggattggc  | gataaaaaac  | aactgctgac  | gccgcgtgc   | gatcagttca  | cccggtgcacc | 3720 |
| gctggataac  | gacattggcg  | taagtgaagc  | gaccgcatt   | gaccctaacg  | cctgggtcga  | 3780 |
| acgctggaaag | gcggcgggccc | attaccaggc  | cgaagcagcg  | ttgttgcagt  | gcacggcaga  | 3840 |
| tacacttgct  | gatgcgggtgc | tgattacgac  | cgctcacgcg  | tggcagcatac | agggggaaaac | 3900 |
| cttattttatc | agccggaaaa  | cctaccggat  | tgatggtagt  | ggtcaaattgg | cgattaccgt  | 3960 |
| tgatgttgaa  | gtggcgagcg  | atacaccgca  | tccggcgcgg  | attggcctga  | actgccagct  | 4020 |
| ggcgcaggta  | gcagagcggg  | taaactggct  | cggatttaggg | ccgcaagaaa  | actatcccga  | 4080 |
| ccgccttact  | gccgcctgtt  | ttgaccgctg  | ggatctgcca  | ttgtcagaca  | tgtatacccc  | 4140 |
| gtacgtcttc  | ccgagcgaaa  | acggtctgcg  | ctgcgggacg  | cgcgaattga  | attatggccc  | 4200 |
| acaccagtgg  | cgcggcact   | tccagttcaa  | catcagccgc  | tacagtcaac  | agcaactgtat | 4260 |
| ggaaaccaggc | catgcaccatc | tgctgcacgc  | ggaagaaggc  | acatggctga  | atatcgacgg  | 4320 |
| tttccatatg  | gggattggtg  | gchgacgactc | ctggagcccg  | tcagttatcgg | cggaattcca  | 4380 |
| gctgagcgcc  | ggtcgctacc  | attaccagtt  | ggtctggtgt  | caaaaaagat  | ctggaggtgg  | 4440 |
| tggcagcagg  | cctggcgcgcg | ccggatcctt  | aattaaacaat | tgaccggtaa  | taataggttag | 4500 |
| ataagtgact  | gattagatgc  | attgatccct  | cgaccaattc  | cggttatttt  | ccaccatatt  | 4560 |
| gccgtctttt  | ggcaatgtga  | gggccccggaa | acctggccct  | gtcttcttga  | cgagcattcc  | 4620 |
| taggggtctt  | tcccctctcg  | ccaaaggaat  | gcaaggtctg  | ttgaatgtcg  | tgaaggaagc  | 4680 |
| agttcctctg  | gaagcttctt  | gaagacaaac  | aacgtctgta  | gcgaccctt   | gcaggcagcg  | 4740 |
| gaacccccc   | cctggcgaca  | ggtgcctctg  | cgccaaaag   | ccacgtgtat  | aagatacacc  | 4800 |
| tgcaaaggcg  | gcacaacccc  | agtgccacgt  | tgtgagttgg  | atagttgtgg  | aaagagtcaa  | 4860 |
| atggctctcc  | tcaagcgtat  | tcaacaaggg  | gctgaaggat  | gcccagaagg  | tacccatttgc | 4920 |
| tatggatct   | gatctggggc  | ctcggtgcac  | atgctttaca  | tgtgttttagt | cgaggttaaa  | 4980 |

On the 27th of May, 1852, the author left New York for the West, in company with his wife and son, and arrived at St. Louis on the 31st.

aaacgtctag gccccccgaa ccacggggac gtggtttcc tttgaaaaac acgatgataa 5040  
taccatgatt gaacaagatg gattcacgc aggttctccg gccgcttggg tggagaggct 5100  
attcggctat gactgggcac aacagacaat cggctgctct gatgccgccc tgccggct 5160  
gtcagcgcag gggcgcccgg ttctttgt caagaccgac ctgtccggtg ccctgaatga 5220  
actgcaggac gaggcagcgc ggctatcgta gctggccacg acggcggtc cttgcgcagc 5280  
tgtgctcgac gtgtcaactg aagcgggaag ggactggctg ctattggcg aagtgcgggg 5340  
gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gatccatca tggctgatgc 5400  
aatgcggcgg ctgcatacgc ttgatccggc tacctgccc ttgcaccacc aagcgaacaca 5460  
tcgcacatcgag cgagcacgta ctggatgga agccggctt gtcgatcagg atgatctgga 5520  
cgaagagcat caggggctcg cgccagccga actgttcgcc aggctcaagg cgccatgcc 5580  
cgacggcgag gatctcgta tgaccatgg cgatgcccgc ttgccaata tcatggtgg 5640  
aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcg accgctatca 5700  
ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg 5760  
cttcctcgta ctttacggta tcgcccgtcc cgattcgacg cgcatcgcc tctatcgcc 5820  
tcttgcacgag ttcttctgag cgggactctg ggttcgcat cgataaaata aaagattta 5880  
tttagtctcc agaaaaaggg gggaatgaaa gaccccaccc gtaggttgg caagctagct 5940  
taagtaacgc cattttgcaa ggcattggaaa aatacataac tgagaataga gaagttcaga 6000  
tcaagggtcag gaacagatgg aacagactgaa tatggccaa acaggatatc tgtggtaagc 6060  
agttcctgcc ccggctcagg gccaagaaca gatggacacag ctgaatatgg gccaaacagg 6120  
atatctgtgg taagcagttc ctgccccggc tcagggccaa gaacagatgg tccccagatg 6180  
cggccatgcc ctcagcagtt tctagagaac catcagatgt ttccagggtg ccccaaggac 6240  
ctgaaatgac cctgtgcctt atttgaacta accaatcagt tcgcttctcg cttctgtcg 6300  
cgccgttctg ctccccgagc tcaataaaag agcccacaac ccctcactcg gggcccaagt 6360  
cctccgattt gactgagtcgc ccgggtaccc gtgtatccaa taaaccctct tgcagttgca 6420  
tccgacttgt ggtctcgctg ttccttgggaa gggtctcctc tgagtgattt actacccgtc 6480  
agcgggggtc tttcattcat gcagcatgta tcaaaattaa tttggttttt tttcttaagt 6540  
atttacatta aatggccata gttgcattaa tgaatcggcc aacgcgcggg gagaggcggt 6600  
ttgcgtatgg gcgctttcc gttcctcgcc tcactgactc gctgcgctcg gtcgttccggc 6660  
tgcggcgacq qqtatcaqct cactcaaaqq cqgtaatacq qttatccaca qaatcaqqqq 6720

|   |      |
|---|------|
| ataaacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg  | 6780 |
| ccgcgttgct ggcgttttc cataggctcc gccccctga cgagcatcac aaaaatcgac     | 6840 |
| gctcaagtca gaggtggcga aacccgacag gactataag ataccaggcg tttccccctg    | 6900 |
| gaagctccct cgtgcgtct cctgttccga .ccctgccgct taccggatac ctgtccgcct   | 6960 |
| ttctcccttc gggaaagcgtg gcgccttctc atagctcacg ctgtaggtat ctcagttcgg  | 7020 |
| tgttaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct  | 7080 |
| gcgccttatac cggttaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac | 7140 |
| tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt   | 7200 |
| tcttgaagtg gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgtc    | 7260 |
| tgctgaagcc agttacccctc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca  | 7320 |
| ccgctggtag cggtggggtt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat   | 7380 |
| ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggAAC gaaaactcac   | 7440 |
| gttaaggat tttggtcatg agattatcaa aaaggatctt cacctagatc ctttgcggc     | 7500 |
| cgcaaataa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc    | 7560 |
| agtgaggcac ctatctcagc gatctgtcta tttcggtcat ccatagttgc ctgactcccc   | 7620 |
| gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata   | 7680 |
| ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg   | 7740 |
| gccgagcgcga gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc  | 7800 |
| cggaaagcta gagtaagtag ttcgcccagt aatagttgc gcaacgttgt tgccattgct    | 7860 |
| acaggcatcg tggtgtcacg ctcgtcggtt ggtatggctt cattcagctc cggttcccaa   | 7920 |
| cgtatcaaggc gagttacatg atccccatg ttgtcaaaa aagcggttag ctccttcgg     | 7980 |
| cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca   | 8040 |
| ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac   | 8100 |
| tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctctg cccggcgtca    | 8160 |
| atacgggata ataccgcgc acatagcaga actttaaaag tgctcatcat tggaaaacgt    | 8220 |
| tcttcggggc gaaaactctc aaggatctt ccgctgttga gatccagttc gatgtAACCC    | 8280 |
| actcgtgcac ccaactgatc ttcagcatct tttacttca ccagcggttc tgggtgagca    | 8340 |
| aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata   | 8400 |
| ctcatactct tccttttca atattattga agcatttatac agggttatttgc tctcatgagc | 8460 |

ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggccgcg cacattc 8518

<210> 5  
<211> 8518  
<212> DNA  
<213> Unknown

<220>  
<223> pICAST ALN.

|   |      |
|---|------|
| <400> 5   |      |
| gacgtccggac ttataccgg tttgtcctat agacaccatt cgtcaaggac gggcccgagt     | 60   |
| cccggttctt gtctaccttgc tcgacttata cccggtttgt cctatagaca ccattcgtca    | 120  |
| aggacggggc cgagtccgg ttcttgcata ccaggggtct acgcccaggc gggagtcgtc      | 180  |
| aaagatctct tggtagtcta caaagggtccc acggggttcc tggactttac tggacacgg     | 240  |
| aataaaacttgc attggtagt caagcgaaga gcgaagacaa gcgcgcgaag acgaggggct    | 300  |
| cgagttattt tctcgggtgt tggggagtga gccccgcgtt caggaggcta actgactcag     | 360  |
| cgggcccattt ggcacatagg ttatttggga gaacgtcaac gttaggctgaa caccagagcg   | 420  |
| acaaggaacc ctcccagagg agactcacta actgatggc agtcgcccc agaaagtaaa       | 480  |
| cccccgagca ggccctagcc ctctggggac gggtcctgg tggctgggtg gtggccctcc      | 540  |
| gttcgaccgg tcgttgaata gacacagaca ggctaacaga tcacagatac tgactaaaat     | 600  |
| acgcggacgc agccatgatc aatcgattga tcgagacata gaccgcctgg gcaccacctt     | 660  |
| gactgctcaa gacttgggg ccggcggtgg gaccctctgc agggtccctg aaaccccccgg     | 720  |
| caaaaacacc gggctggact cttccctca gctacacctt aggctggggc agtccctatac     | 780  |
| accaagacca tcctctgctc ttggatttttgc tcaaggccgg aggcagactt aaaaacgaaa   | 840  |
| gccaaacctt ggcttcggcg cgccagaacag acgacgtcgt agcaagacac aacagagaca    | 900  |
| gactgacaca aagacataaa cagacttttta atcccggtct gacaatggtg agggattca     | 960  |
| aactggaaatc cattgacctt tctacagctc gccgagcggag tttgggtcag ccatctacag   | 1020 |
| ttcttctctg caacccaaatg gaagacgaga cgtcttaccg gttggaaatt gcagcctacc    | 1080 |
| ggcgctctgc cgtggaaatt ggctctggag tagtgggtcc aattctagtt ccagaaaagt     | 1140 |
| ggaccggggc tacctgtggg tctggtccag gggatgttagc actggaccct tcggaaaccga   | 1200 |
| aaactggggg gagggaccca gttcgggaaa catgtggat tcggaggccgg aggagaagga     | 1260 |
| ggttaggcggg gcagagaggg ggaacttggaa ggagcaagct gggccggagc taggaggaa    | 1320 |
| ataggtcggtt agtggagaaatg agatccgcgg ccggcgagat cggtaatttgc tgctgagtga | 1380 |

|            |             |            |             |             |             |      |
|------------|-------------|------------|-------------|-------------|-------------|------|
| tatcccgcta | agcttgtggt  | acgtggtagt | agtagtagtg  | cagctgatat  | ttctactcct  | 1440 |
| ggagctctac | ccgcaactaat | gcctaagtga | ccggcagcac  | cgggcgtggc  | tagcgggaag  | 1500 |
| ggttgtcaat | gcgtcggact  | taccgcttac | cgcgaaacgg  | accaaaggcc  | gtggtcttcg  | 1560 |
| ccacggcctt | tcgaccgacc  | tcacgctaga | aggactccgg  | ctatgacagc  | agcagggag   | 1620 |
| tttgaccgtc | tacgtgccaa  | tgctacgcgg | gtagatgtgg  | ttgcactgga  | taggtaatg   | 1680 |
| ccagttaggc | ggcaaacaag  | ggtgccttt  | aggctgccc   | acaatgagcg  | agtgtaaatt  | 1740 |
| acaactactt | tcgaccgatg  | tccttcgg   | ctgcgcttaa  | taaaaaactac | cgcaattgag  | 1800 |
| ccgcaaagta | gacaccacgt  | tgcccgac   | ccagccatg   | ccggcctgt   | cagcaaacgg  | 1860 |
| cagacttaaa | ctggactcgc  | gtaaaaatgc | gcggccttt   | ttggcggagc  | gccactacca  | 1920 |
| cgacgcgacc | tcactgccgt  | caatagacct | tctagtccta  | tacaccgcct  | actcgccgta  | 1980 |
| aaaggcactg | cagagcaacg  | acgtatttg  | ctgatgtgtt  | tagtcgctaa  | aggtacaacg  | 2040 |
| gtgagcgaaa | ttactactaa  | agtcggcg   | acatgacctc  | cgacttcaag  | tctacacgc   | 2100 |
| gctcaacgca | ctgatggatg  | cccattgtca | aagaataacc  | gtcccacttt  | gcgtccagcg  | 2160 |
| gtcgccgtgg | cgcggaaagc  | cgcacttta  | atagctactc  | gcaccaccaa  | tacggctagc  | 2220 |
| gcagtgtat  | gcagacttgc  | agctttggg  | cttgacacc   | tcgcggctt   | agggcttaga  | 2280 |
| gatagcacgc | caccaacttg  | acgtgtggcg | gctgcgtgc   | gactaacttc  | gtcttcggac  | 2340 |
| gctacagcca | aaggcgctcc  | acgcctaact | tttaccagac  | gacgacgact  | tgccgttcgg  | 2400 |
| caacgactaa | gctccgcaat  | tggcagtgc  | cgtagtagga  | gacgtaccag  | tccagtacct  | 2460 |
| actcgtctgc | taccacgtcc  | tataggacga | ctactcgtc   | ttgttgaat   | tgccggcacgc | 2520 |
| gacaagcgta | ataggcttgg  | taggcacac  | catgtgcgac  | acgctggcga  | tgccggacat  | 2580 |
| acaccaccta | cttcggttat  | aactttgggt | gccgtaccac  | ggttacttag  | cagactggct  | 2640 |
| actaggcgcg | accgatggcc  | gctactcgct | tgcgcatatgc | gcttaccacg  | tcgcgttagc  | 2700 |
| attagtgggc | tcacactagt  | agaccagcg  | ccccttactt  | agtccggcgc  | cgcgattagt  | 2760 |
| gctgcgcgac | atagcgacct  | agtttagaca | gctaggaagg  | gcgggcccacg | tcataactcc  | 2820 |
| gccgcctcgg | ctgtggtgcc  | ggtggctata | ataaacgggc  | tacatgcgac  | cgcacctact  | 2880 |
| tctggtcggg | aagggcccac  | acggcttac  | caggtagttt  | tttaccgaaa  | gcgtatggacc | 2940 |
| tctctgcgcg | ggcgactagg  | aaacgcttat | gcgggtgcgc  | tacccattgt  | cagaaccgcc  | 3000 |
| aaagcgattt | atgaccgtcc  | gcaaagcagt | catagggca   | aatgtcccgc  | cgaagcagac  | 3060 |
| cctgacccac | ctagtcagcg  | actaatttat | actactttg   | ccgttggca   | ccagccgaat  | 3120 |

|                       |             |             |             |            |            |      |
|-----------------------|-------------|-------------|-------------|------------|------------|------|
| gccgccacta aaaccgctat | gcggcttgct  | agcggtaag   | acataacttc  | cagaccagaa | 3180       |      |
| acggctggcg            | tgccggcttag | gtcgcgactg  | cttcggttt   | gtggcgctcg | tcaaaaaggt | 3240 |
| caaggcaaat            | aggcccgtt   | ggtagcttca  | ctggcgctt   | atggacaagg | cagtatcgct | 3300 |
| attgctcgag            | gacgtgacct  | accaccgcga  | cctaccattc  | ggcgaccgtt | cgccacttca | 3360 |
| cggagaccta            | cagcgaggtg  | ttccatttgc  | caactaactt  | gacggacttg | atggcgctgg | 3420 |
| cctctcgcg             | cccgttgaga  | ccgagtgtca  | tgcgcacac   | gttggcttgc | gctggcgtag | 3480 |
| cagtcttcgg            | cccgtgttagt | cgcggaccgt  | cgtcaccgca  | gaccgcctt  | tggagtcaca | 3540 |
| ctgcgagggg            | cggcgacagg  | tgcggtaggg  | cgtagactgg  | tggcgcttt  | acctaaaaac | 3600 |
| gtagctcgac            | ccattattcg  | caaccgttaa  | attggcggtc  | agtccgaaag | aaagtgtcta | 3660 |
| cacctaaccg            | ctatttttg   | ttgacgactg  | cggcgacgca  | ctagtcaga  | gggcacgtgg | 3720 |
| cgacctattg            | ctgttaaccgc | attcacttcg  | ctggcgtaa   | ctgggattgc | ggaccaggct | 3780 |
| tgcgacccccc           | cgcgcgg     | taatggtccg  | gttcgtcgc   | aacaacgtca | cgtgccgtct | 3840 |
| atgtgaacga            | ctacgccacg  | actaatgctg  | gcgagtgcgc  | accgtcgtag | tcccctttg  | 3900 |
| gaataaaatag           | tccgccttt   | ggatggccta  | actaccatca  | ccagtttacc | gctaatggca | 3960 |
| actacaactt            | caccgctcgc  | tatgtggcgt  | aggccgcgc   | taaccggact | tgcggcgtca | 4020 |
| ccgcgtccat            | cgtctcgccc  | atttgaccga  | gcctaattccc | ggcgttcttt | tgatagggct | 4080 |
| ggcggatga             | cggcgacaa   | aactggcgcac | cctagacggt  | aacagtctgt | acatatgggg | 4140 |
| catgcagaag            | ggctcgctt   | tgcacacgc   | gacgcctgc   | gcgcctaact | taataccggg | 4200 |
| tgtggtcacc            | gcgcgcgtga  | aggtcaagtt  | gtagtcggcg  | atgtcagttg | tcgttgacta | 4260 |
| ccttggtcg             | gtagcggtag  | acgacgtgcg  | cttcttcgg   | tgtaccgact | tatagtgcgc | 4320 |
| aaaggtatac            | ccctaaccac  | cgctgctgag  | gacctcgggc  | agtcatagcc | gccttaaggt | 4380 |
| cgactcgcgg            | ccagcgatgg  | taatggtcaa  | ccagaccaca  | gtttttctta | gacctccacc | 4440 |
| accgtcgcc             | ggaaccgcgc  | ggcctaggaa  | ttaattgtta  | actggccatt | attatccatc | 4500 |
| tattcactga            | ctaattctacg | taactaggga  | gctggtaag   | gccaataaaa | ggtgtataaa | 4560 |
| cggcagaaaa            | ccgttacact  | cccgccctt   | tggaccggga  | cagaagaact | gctcgtaagg | 4620 |
| atccccagaa            | aggggagagc  | ggtttcctta  | cgttccagac  | aacttacagc | acttccttcg | 4680 |
| tcaaggagac            | cttcgaagaa  | cttctgtttt  | ttgcagacat  | cgctggaaa  | cgtccgtcgc | 4740 |
| cttgggggggt           | ggaccgctgt  | ccacggagac  | gccggtttc   | ggtgcacata | ttctatgtgg | 4800 |
| acgtttccgc            | cgtgttgggg  | tcacggtgca  | acactcaacc  | tatcaacacc | tttctcagtt | 4860 |

|   |      |
|---|------|
| taccgagagg agttcgata agttgttccc cgacttcata cgggtcttcc atggggtaac    | 4920 |
| ataccctaga ctagaccccg gagccacgtg tacgaaatgt acacaaatca gctccaattt   | 4980 |
| tttgcagatc cggggggctt ggtgcccctg cacaaaagg aaacttttg tgctactatt     | 5040 |
| atggtaactaa cttgttctac ctaacgtgcg tccaagaggc cggcgaaccc acctctccga  | 5100 |
| taagccata ctgaccctgt ttgtctgtta gccgacgaga ctacggcggc acaaggccga    | 5160 |
| cagtcgcgtc cccgcgggccc aagaaaaaca gttctggctg gacaggccac gggacttact  | 5220 |
| tgacgtcctg ctccgtcgcg ccgatagcac cgaccggc tgccgcgaag gaacgcgtcg     | 5280 |
| acacgagctg caacagtgc ttcgccccttc cctgaccgac gataaccgc ttcacggccc    | 5340 |
| cgtcctagag gacagtagag tggAACGAGG acggctctt cataggtgt accgactacg     | 5400 |
| ttacgcccgc gacgtatgcg aactaggccg atggacgggt aagctggtgg ttgcctttgt   | 5460 |
| agcttagctc gctcgtcat gaggctacct tcggccagaa cagctagtc tactagacct     | 5520 |
| gcttctcgta gtccccgagc gcgggtcggct tgacaagcgg tccgagttcc gcgcgtacgg  | 5580 |
| gctgccgtc ctagagcagc actgggtacc gctacggacg aacggcttat agtaccacct    | 5640 |
| tttaccggcg aaaagaccta agtagctgac accggccgac ccacaccgc tggcgatagt    | 5700 |
| cctgtatcgc aaccgatggg cactataacg acttctcgaa ccgcccgtt cccgactggc    | 5760 |
| gaaggagcac gaaatgccat agcggcgagg gctaagcgtc gcgtagcgga agatagcgga   | 5820 |
| agaactgctc aagaagactc gccctgagac cccaaagcgt a gctattttat tttctaaaat | 5880 |
| aaatcagagg tcttttccc cccttacttt ctgggggtgg catccaaacc gttcgatcga    | 5940 |
| attcattcgc gtaaaacgtt ccgtacctt ttatgtattt actcttatct cttcaagtct    | 6000 |
| agttccagtc cttgttctacc ttgtcgactt ataccgggtt tgtcctatac acaccattcg  | 6060 |
| tcaaggacgg ggccgagtcc cggttcttgc ctacccgtc gacttatacc cggttgc       | 6120 |
| tatagacacc attcgtcaag gacggggccg agtcccggtt cttgttctacc aggggtctac  | 6180 |
| gccaggtcgg gagtcgtcaa agatctctt gtagtctaca aaggtcccac ggggtccctg    | 6240 |
| gactttactg ggacacggaa taaacttgat tggtagtca agcgaagagc gaagacaagc    | 6300 |
| gccccggcgg agttatcc tcgggtgtt gggagtgagc cccgcggta                  | 6360 |
| ggaggctaact gactcagcg ggcccatggg cacataggtt atttgggaga acgtcaacgt   | 6420 |
| aggctgaaca ccagagcgac aaggaaccct cccagaggag actcaactaac tgatggcag   | 6480 |
| tcgccccccag aaagtaagta cgtcgtacat agtttaattt aaaccaaaaa aaagaattca  | 6540 |
| taaatgtaat ttaccggat caacgttaattt acttagccgg ttgcgcgccttctccgc      | 6600 |

Sequence

|             |             |             |            |            |             |      |
|-------------|-------------|-------------|------------|------------|-------------|------|
| aacgcataac  | cgcgagaagg  | cgaaggagcg  | agtgactgag | cgacgcgagc | cagcaagccg  | 6660 |
| acgcccgtcg  | ccatagtcga  | gtgagttcc   | gccattatgc | caataggtgt | cttagtcccc  | 6720 |
| tattgcgtcc  | tttcttgtac  | actcgtttc   | cggcggttt  | ccggccttg  | gcattttcc   | 6780 |
| ggcgcaacga  | ccgcaaaaag  | gtatccgagg  | cggggggact | gctcgtagtg | tttttagctg  | 6840 |
| cgagttcagt  | ctccaccgct  | ttgggctgtc  | ctgatatttc | tatggtccgc | aaagggggac  | 6900 |
| cttcgaggga  | gcacgcgaga  | ggacaaggct  | gggacggcga | atggcctatg | gacaggcgg   | 6960 |
| aagagggaaag | cccttcgcac  | cgcgaaagag  | tatcgagtgc | gacatccata | gagtcaagcc  | 7020 |
| acatccagca  | agcgaggttc  | gacccgacac  | acgtgcttg  | ggggcaagtc | gggctggcga  | 7080 |
| cgcggaatag  | gccattgata  | gcagaactca  | ggttggcca  | ttctgtgctg | aatagcggt   | 7140 |
| accgtcgtcg  | gtgaccattg  | tcctaattcgt | ctcgctccat | acatccgcca | cgatgtctca  | 7200 |
| agaacttcac  | caccggattt  | atgccgatgt  | gatctcttg  | tcataaacca | tagacgcgag  | 7260 |
| acgacttcgg  | tcaatggaag  | ccttttctc   | aaccatcgag | aactaggccg | tttgggggt   | 7320 |
| ggcgaccatc  | gccaccaaaa  | aaacaaacgt  | tcgtcgtcta | atgcgcgtct | tttttccta   | 7380 |
| gagttcttct  | aggaaactag  | aaaagatgcc  | ccagactgcg | agtcacctt  | ctttgagtg   | 7440 |
| caattcccta  | aaaccagtac  | tctaatagtt  | ttccctagaa | gtggatctag | gaaaacgccc  | 7500 |
| gcgtttagtt  | agatttcata  | tatactcatt  | tgaaccagac | tgtcaatgg  | tacgaattag  | 7560 |
| tcactccgtg  | gatagagtcg  | ctagacagat  | aaagcaagta | ggtatcaacg | gactgaggg   | 7620 |
| cagcacatct  | attgtatgcta | tgccctcccg  | aatggtagac | cggggtcacg | acgttactat  | 7680 |
| ggcgctctgg  | gtgcgagtgg  | ccgaggtcta  | aatagtcgtt | atttggtcgg | tcggccttcc  | 7740 |
| cggctcgcgt  | cttcaccagg  | acgttgaat   | aggcggaggt | aggtcagata | attaaacaacg | 7800 |
| gcccttcgat  | ctcattccatc | aagcggtcaa  | ttatcaaacg | cgttgcaaca | acggtaacga  | 7860 |
| tgtccgtagc  | accacagtgc  | gagcagcaaa  | ccataccgaa | gtaagtcgag | gccaagggtt  | 7920 |
| gctagttccg  | ctcaatgtac  | taggggtac   | aacacgttt  | ttcgccaatc | gaggaagcca  | 7980 |
| ggaggctagc  | aacagtcttc  | attcaaccgg  | cgtcacaata | gtgagtagca | ataccgtcgt  | 8040 |
| gacgtattaa  | gagaatgaca  | gtacggtagg  | cattctacga | aaagacactg | accactcatg  | 8100 |
| agttggttca  | gtaagactct  | tatcacatac  | gccgctggct | caacgagaac | gggcccgcagt | 8160 |
| tatgccctat  | tatggcgcgg  | tgtatcgtct  | tgaattttc  | acgagtagta | acctttgca   | 8220 |
| agaagccccc  | cttttgagag  | ttcctagaat  | ggcgacaact | ctaggtcaag | ctacattggg  | 8280 |
| tgagcacgtg  | ggttgactag  | aagtgcgtaga | aaatgaaagt | ggtcgcaaag | acccactcgt  | 8340 |

|   |      |
|---|------|
| ttttgtcctt ccgttttacg gcgtttttc ctttattccc gctgtgcctt tacaacttat    | 8400 |
| gagttatgaga aggaaaaagt tataataact tcgtaaatag tcccaataac agagtaactcg | 8460 |
| cctatgtata aacttacata aatctttta tttgtttatc cccaggcgc gtgtaaag       | 8518 |

<210> 6  
<211> 8175  
<212> DNA  
<213> Unknown

<220>  
<223> pICAST OMC.

|  |      |
|--|------|
| <400> 6<br>ctgcagcctg aatatggcc aaacaggata tctgtggtaa gcagttcctg cccggctca | 60   |
| gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt          | 120  |
| tcctgccccg gctcaggcc aagaacagat ggtccccaga tgccgtccag ccctcagcag           | 180  |
| tttctagaga accatcagat gtttccaggg tgcccaagg acctgaaatg accctgtgcc           | 240  |
| ttatttgaac taaccaatca gttcgcttct cgcttctgtt cgccgcgttc tgctccccga          | 300  |
| gctcaataaa agagcccaca acccctcaact cggggcgcca gtcctccgat tgactgagtc         | 360  |
| gccgggtac ccgtgtatcc aataaaccct cttgcagttt catccgactt gtggtctcgc           | 420  |
| tgttccttgg gaggytctcc tctgagtgtat tgactacccg tcagcggggg tcttcattt          | 480  |
| gggggctcgt ccgggatcgg gagacccctg cccagggacc accgaccac caccggagg            | 540  |
| caagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgattta           | 600  |
| tgcgcctgctg tcggtaact agctgttat ctggcggacc cgtggtgaa                       | 660  |
| ctgacgagtt ctgaacaccc ggccgcaacc ctggagacg tcccaggac tttgggggccc           | 720  |
| gtttttgtgg cccgacctga ggaaggaggt cgatgtggaa tccgaccctcg tcaggatatg         | 780  |
| tggttctgggt aggagacgag aacctaaaac agttcccgcc tccgtctgaa tttttgcttt         | 840  |
| cggtttggaa ccgaagccgc gcgtcttgc tgctgcagca tcgttctgtt ttgtctctgt           | 900  |
| ctgactgtgt ttctgtattt gtctgaaaat tagggccaga ctgttaccac tcccttaagt          | 960  |
| ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccagtc ggtagatgtc          | 1020 |
| aagaagagac gttgggttac ctctgctct gcagaatggc caacctttaa cgtcggatgg           | 1080 |
| ccgcgagacg gcacctttaa ccgagacctc atcaccagg ttaagatcaa ggtctttca            | 1140 |
| cctggccgc atggacaccc agaccaggtc ccctacatcg tgacctggga agccttggct           | 1200 |
| tttggacccccc ctccctgggt caagcccttt gtacacccta agcctccgc tccttccct          | 1260 |

|   |      |
|---|------|
| ccatccgccc cgtctctccc ccttgaacct cctcggtcg acccgcctcg atcctccctt      | 1320 |
| tatccagccc tcactccttc tctaggcgcc ggccgctcta gcccattaat acgactca       | 1380 |
| atagggcgat tcgaatcagg ccttggcgcg ccggatcctt aattaagcgc aattgggagg     | 1440 |
| tggcggtac ctcgagatgg gcgtgattac ggattcactg gccgtcggtt tacaacgtcg      | 1500 |
| tgactggaa aaccctggcg ttacccaact taatcgctt gcagcacatc ccccttcgc        | 1560 |
| cagctggcgt aatagcgaag aggccgcac cgatcgccct tcccaacagt tacgcagcct      | 1620 |
| gaatggcgaa tggcgcttg cctggttcc ggcaccagaa gcggtgccgg aaagctggct       | 1680 |
| ggagtgcgat cttcctgagg cogatactgt cgtcgcccc tcaaactggc agatgcacgg      | 1740 |
| ttacgatgac cccatctaca ccaacgtgac ctatcccatt acggtaatc cgccgttgtt      | 1800 |
| tcccacggag aatccgacgg gttgttactc gtcacatatt aatgttgatg aaagctggct     | 1860 |
| acaggaaggc cagacgcgaa ttattttga tggcgtaac tcggcggttc atctgtggtg       | 1920 |
| caacgggcgc tgggtcggtt acggccagga cagtcgtttc ccgtctgaat ttgacctgag     | 1980 |
| cgcattttta cgcgccccggag aaaaccgcct cgcggtgatg gtgctgcgt ggagtgcacgg   | 2040 |
| cagttatctg gaagatcagg atatgtggcg gatgagcggc atttccgtt acgtctcggtt     | 2100 |
| gctgcataaaa ccgactacac aaatcagcga tttccatgtt gccactcgct ttaatgtga     | 2160 |
| tttcagccgc gctgtactgg aggctgaagt tcagatgtgc ggcgagttgc gtgactacct     | 2220 |
| acgggtaaca gtttctttat ggcagggtga aacgcaggc gccagcggca ccgcgcctt       | 2280 |
| cggcggtgaa attatcgatg agcgtggtgg ttatgccat cgcgtcacac tacgtctgaa      | 2340 |
| cgtcgaaaac ccgaaaactgt ggagcggccga aatcccgaat ctctatcggtt ccgtgggttga | 2400 |
| actgcacacc gccgacggca cgctgattga agcagaagcc tgcgatgtcg gtttccgcga     | 2460 |
| ggtgccgatt gaaaatggtc tgctgctgtt gaacggcaag ccgttgctga ttcgaggcgt     | 2520 |
| taaccgtcac gagcatcatc ctctgcatgg tcaggtcatg gatgagcaga cgatggcgt      | 2580 |
| ggatatacctg ctgatgaagc agaacaactt taacgccgtg cgctgttcgc attatccgaa    | 2640 |
| ccatccgctg tggcacacgc tgtgcgaccg ctacggcctg tatgtggtgg atgaagccaa     | 2700 |
| tattgaaacc cacggcatgg tgccaatgaa tcgtctgacc gatgatccgc gctggctacc     | 2760 |
| ggcgatgagc gaacgcgtaa cgcgaatggt gcagcgcgtat cgtaatcacc cgagtgtgtat   | 2820 |
| catctggtgc ctggggatgt aatcaggcca cggcgctaat cacgacgcgc tgcgtatcggt    | 2880 |
| gatcaaatttct gtcgatcctt cccgccccgt gcagttatgaa ggcggcggag ccgacaccac  | 2940 |
| ggccaccgat attatttgcc cgatgtacgc gcgcgtggat gaagaccagc ccttccggc      | 3000 |

tggccgaaa tggccatca aaaaatggct ttcgctacct ggagagacgc gcccgtgat 3060  
 ccttgcgaa tacgcccacg cgatggtaa cagtctggc gtttcgcta aatactggca 3120  
 ggcgttcgt cagtatcccc gttacaggg cggcttcgtc tggactggg tggatcagtc 3180  
 gctgattaaa tatgatgaaa acggcaaccc gtggcggct tacggcggtg attttggcga 3240  
 tacgcccgaac gatcgccagt tctgtatgaa cggctcggtc tttgccgacc gcacgcccga 3300  
 tccagcgctg acggaagcaa aacaccagca gcagttttc cagttccgtt tatccggca 3360  
 aaccatcgaa gtgaccagcg aatacctgtt ccgtcatagc gataacgagc tcctgcactg 3420  
 gatggtggcg ctggatggta agccgctggc aagcggtgaa gtgcctctgg atgtcgctcc 3480  
 acaaggtaaa cagttgattt aactgcctga actaccgcag ccggagagcg ccgggcaact 3540  
 ctggctcaca gtacgcgttag tgcaaccgaa cgcgaccgca tggtcagaag ccgggcacat 3600  
 cagcgcctgg cagcagtggc gtctggcgga aaacctcagt gtgacgctcc ccggccgtc 3660  
 ccacgcccattt ccgcatttgcgaa ccaccagcgaa aatggatttt tgcatcgagc tggtaataa 3720  
 gcgttggcaa tttaaccgccc agtcaggctt tctttcacag atgtggattt gcgataaaaa 3780  
 acaactgctg acgcccgtgc gcgcatttttgcgaa caccgtgtc gatagatctg aacagaaact 3840  
 catttccgaa gaagacctag tcgaccatca tcattatcat caccgttaat aataggtaga 3900  
 taagtgactg attagatgca ttgcacttag atccctcgac caattccgtt tattttccac 3960  
 catattgccg tctttggca atgtgagggc ccggaaacctt ggccctgtct tcttgacgag 4020  
 cattcctagg ggtctttccc ctctcgccaa aggaatgcaaa ggtctgttga atgtcgtaa 4080  
 ggaagcagtt cctctggaag cttcttgaag acaaacaacg tctgtacgca cccttgcag 4140  
 gcagcggAACccccacctg gcgcacagggtg cctctgcggc caaaagccac gtgtataaga 4200  
 tacacctgca aaggcggcac aacccctgtt ccacgttgcg agttggatag ttgtggaaag 4260  
 agtcaaattttt ctctcctcaa gcgttatttcaa caaggggtcg aaggatgccc agaaggatacc 4320  
 ccattgtatg ggatctgatc tggggcctcg gtgcacatgc tttacatgtt ttttagtcgag 4380  
 gttaaaaaac gtcttaggccc cccgaaccac ggggacgtgg ttttcctttt aaaaacacgaa 4440  
 tgataataacc atgaaaaaagc ctgaactcac cgccgcgtct gtcgagaatgt ttctgtatcgaa 4500  
 aaagttcgac agcgtctccg acctgtatgca gctctcgag ggcgaagaat ctcgtgtttt 4560  
 cagcttcgtatgtt gtaggaggcgtt gttttttttt aatagctgcg ccgtatggttt 4620  
 ctacaaagat cgttatgtttt atcggcactt tgcatcgcc ggcgtcccgaa ttccggaaatgtt 4680  
 gcttgacatt gggaaattta gcgagagcctt gacattttgc atctcccgcc gtgcacagggtt 4740

|              |              |              |             |             |              |      |
|--------------|--------------|--------------|-------------|-------------|--------------|------|
| tgtcacgttg   | caagacacctgc | ctgaaaccga   | actgcccgt   | gttctgcagc  | cggtcgcgga   | 4800 |
| ggccatggat   | gcgatcgctg   | cggccgatct   | tagccagacg  | agcggggtcg  | gcccattcgg   | 4860 |
| accgcaagga   | atcggtcaat   | acactacatg   | gcgtgatttc  | atatgcgcga  | ttgctgatcc   | 4920 |
| ccatgtgtat   | cactggcaaa   | ctgtgatgga   | cgacaccgtc  | agtgcgtccg  | tcgcgcaggc   | 4980 |
| tctcgatgag   | ctgatgctt    | gggcccggagga | ctgccccgaa  | gtccggcacc  | tcgtgcacgc   | 5040 |
| ggatttcggc   | tccaacaatg   | tcctgacgga   | caatggccgc  | ataacagcgg  | tcattgactg   | 5100 |
| gagcgaggcg   | atgttcgggg   | attcccaata   | cgaggtcgcc  | aacatcttct  | tctggaggcc   | 5160 |
| gtgggtggct   | tgtatggagc   | agcagacgctg  | ctacttcgag  | cgaggcattc  | cggagcttgc   | 5220 |
| aggatcgccg   | cggctccggg   | cgtatatgct   | ccgcattggt  | cttgcaccaac | tctatcagag   | 5280 |
| cttgggttgcac | ggcaatttcg   | atgatgcagc   | ttgggcgcag  | ggtcgatgcg  | acgcaatcgt   | 5340 |
| ccgatccgga   | gccgggactg   | tcgggcgtac   | acaaatcgcc  | cgcagaagcg  | cggccgtctg   | 5400 |
| gaccgatggc   | tgtgtagaag   | tactcgccga   | tagtgaaac   | cgacgccccca | gcactcggtcc  | 5460 |
| gagggcaaag   | gaatagagta   | gatgccgacc   | gggatctatc  | gataaaataa  | aagattttat   | 5520 |
| ttagtctcca   | aaaaaaagggg  | ggaatgaaag   | accccacctg  | taggtttggc  | aagctagctt   | 5580 |
| aagtaacgcc   | attttgcag    | gcatggaaaaa  | atacataact  | gagaatagag  | aagttcagat   | 5640 |
| caaggtcagg   | aacagatgga   | acagctgaat   | atgggcaaaa  | caggatatct  | gtggtaagca   | 5700 |
| gttcctgccc   | cggctcaggg   | ccaagaacag   | atggaacacgc | tgaatatggg  | ccaaacagga   | 5760 |
| tatctgtggt   | aagcagttcc   | tgccccggct   | cagggccaag  | aacagatggt  | ccccagatgc   | 5820 |
| ggtccagccc   | tcagcagttt   | ctagagaacc   | atcagatgtt  | tccagggtgc  | cccaaggacc   | 5880 |
| tgaaatgacc   | ctgtgcctta   | tttgaactaa   | ccaatcagtt  | cgcttctcgc  | ttctgttcgc   | 5940 |
| gcgcattctgc  | tccccgagct   | caataaaaga   | gcccacaacc  | cctcactcgg  | ggcgccagtc   | 6000 |
| ctccgattga   | ctgagtcgcc   | cgggtaccccg  | tgtatccaat  | aaaccctctt  | gcagttgcatt  | 6060 |
| ccgacttgtg   | gtctcgctgt   | tccttggag    | ggtctccctct | gagtgattga  | ctaccggcgtca | 6120 |
| gcgggggtct   | ttcattcatg   | cagcatgtat   | caaaattaat  | ttggttttt   | ttcttaagta   | 6180 |
| tttacattaa   | atggccatag   | ttgcattaaat  | gaatcgccca  | acgcgcgggg  | agaggcgggtt  | 6240 |
| tgcgtattgg   | cgctcttccg   | cttcctcgct   | cactgactcg  | ctgcgcgtcg  | tcgttcggct   | 6300 |
| gcggcgagcg   | gtatcagctc   | actcaaaggc   | ggtataacgg  | ttatccacag  | aatcagggga   | 6360 |
| taacgcagga   | aagaacatgt   | gagcaaagg    | ccagcaaag   | gccaggaacc  | gtaaaaaggc   | 6420 |
| cgcgttgcgtg  | gcgttttcc    | ataggctccg   | ccccctgac   | gagcatcaca  | aaaatcgacg   | 6480 |

|            |            |            |            |             |             |      |
|------------|------------|------------|------------|-------------|-------------|------|
| ctcaagtca  | aggtggcgaa | acccgacagg | actataaaga | taccaggcgt  | ttccccctgg  | 6540 |
| aagctccctc | gtgcgc     | ctgttccgac | cctgccgctt | accggatacc  | tgtccgc     | 6600 |
| tctcccttcg | ggaagcgtgg | cgcttctca  | tagctcacgc | tgttaggtatc | tcagttcggt  | 6660 |
| gtaggtcg   | ttcgcttca  | ggggctgtgt | gcacgaaccc | cccgttc     | ccgaccgctg  | 6720 |
| cgccttatcc | ggtaactatc | gtcttgagtc | caaccggta  | agacacgact  | tatcgccact  | 6780 |
| ggcagcagcc | actggtaaca | ggattagcag | agcgaggtat | gtaggcggt   | ctacagagtt  | 6840 |
| cttgaagtgg | tggccta    | actac      | tagaagaaca | gtatttggta  | tctgcgctct  | 6900 |
| gctgaagcca | gttac      | ttcg       | aaaaagagt  | tggtagctct  | tgatccggca  | 6960 |
| cgctggtagc | ggtggtttt  | ttgttgc    | gcagcagatt | acgcgcagaa  | aaaaaggatc  | 7020 |
| tcaagaagat | cctt       | gatct      | tttctacgg  | gtctgacgct  | cagtggaaacg | 7080 |
| ttaagggatt | ttgg       | catga      | gattatcaa  | aaggatcttc  | acctagatcc  | 7140 |
| aaaatgaagt | ttgcggccgc | aaatcaatct | aaagtatata | ttagttaact  | tggtctgaca  | 7200 |
| gttaccaatg | cttaatc    | agt        | gaggcaccta | tctcagcgt   | ctgtctattt  | 7260 |
| tagttgcctg | actccccgtc | gtgt       | tagataaa   | ctacgatacg  | ggagggctt   | 7320 |
| ccagtgc    | aatgataacc | cgagacccac | gctcacccgc | tccagattt   | tcaataa     | 7380 |
| accagcc    | cggaagg    | ggcc       | gagcgcagaa | gtggcctgc   | aactttatcc  | 7440 |
| agtctattaa | ttgttgc    | gg         | gaagctagag | taagtagttc  | gccagtt     | 7500 |
| acgtt      | gttgc      | cattgctaca | ggcatcg    | tgtcacg     | gtcg        | 7560 |
| tcag       | ctcc       | aaacga     | tcaaggc    | ttacatgatc  | ccccatgtt   | 7620 |
| cgg        | ttcg       | gtctc      | ccgatcg    | tttgc       | tgcaaaaa    | 7680 |
| tcatg      | gttgc      | tttgc      | cataatt    | ttactgt     | ccatccg     | 7740 |
| ctgtg      | actgg      | ttagtactca | accaagt    | tctgagaata  | gtgtatgc    | 7800 |
| gcttt      | ggttgc     | tttgc      | ggcgtcaata | cgggataata  | ccgc        | 7860 |
| tcatcattgg | aaaacgtt   | tcggggcg   | aactctcaag | gatcttac    | ctgtt       | 7920 |
| ccagttcgat | gtaa       | ccact      | cgtgcaccc  | actgatctc   | agcat       | 7980 |
| gcgtt      | ctgg       | gtgagcaaaa | acaggaaggc | aaaatgc     | ccgc        | 8040 |
| cacggaaatg | ttgaata    | actc       | atactcttcc | ttttcaata   | ttattga     | 8100 |
| gttattgtct | catgagcg   | ga         | tacatattt  | aatgtattt   | aaaaataaa   | 8160 |
| ttccgcgcac | atttc      |            |            |             |             | 8175 |

<210> 7  
<211> 8175  
<212> DNA  
<213> Unknown

<220>  
<223> pICAST OMC.

<400> 7  
gacgtcgac ttatacccg tttgtcctat agacaccatt cgtcaaggac ggggccgagt  
cccggttctt gtctacccgg tcgacttata cccgggttgc cctatagaca ccattcgta  
aggacggggc cgagtcggc ttcttgcata ccaggggtct acgcccggc gggagtcgtc  
aaagatctct tggtagtcta caaagggtccc acggggttcc tggactttac tggacacgg  
aataaacttg attggtagt caagcgaaga gcaagacaa ggcgcgcgaag acgagggct  
cgagttattt tctcggtgt tggggagtga gccccgcgt caggaggcta actgactcag  
cgggcccatg ggcacatagg ttatttggga gaacgtcaac gtaggctgaa caccagagcg  
acaaggaacc ctcccagagg agactcacta actgatggc agtcgcggc agaaagtaaa  
cccccgagca ggccttagcc ctctgggac gggccctgg tggctgggtg gtggccctcc  
gttcgaccgg tcgttgaata gacacagaca ggctaacaga tcacagatac tgactaaaat  
acgcggacgc agccatgatc aatcgattga tcgagacata gaccgcctgg gcaccaccc  
gactgctcaa gacttgggg ccggcggtgg gaccctctgc agggccctg aaaccccccgg  
caaaaacacc gggctggact cttccctca gctacaccc ttggcttgc agtcctatac  
accaagacca tcctctgctc ttggattttgc tcaaggcgccg aggcagactt aaaaacgaaa  
gccaacccctt ggcttcggcg cgcaacacag acgacgtcgt agcaagacac aacagagaca  
gactgacaca aagacataaa cagacttttgc atcccggtct gacaatggtg agggaaattca  
aactggaaatc cattgaccc ttacagctc gcccggcggag tggtggtcag ccatctacag  
ttcttctctg caacccaaatg gaagacgaga cgttttaccg gttggaaattt gcagcctacc  
ggcgctctgc cgtggaaattt ggcttggag tagtgggtcc aattcttagtt ccagaaaaatg  
ggaccggcg tacctgtggg tctggccag gggatgtacg actggaccct tcggaaaccga  
aaactggggg gagggacccca gttcggggaaa catgtggat tcggaggcg aggagaagga  
ggttaggcgg gcagagaggg ggaacttggaa ggagcaagct gggccggc taggaggaa  
ataggtcgaaatc agtgaggaatc agatcccgccg ccggcgagat cggtaattt tgctgagtt  
tatcccgctca aqcttagtcc qqaaccggcgc ggcctaggaa ttaattcgcc ttaaccctcc  
1440

© 2023 The Board of Trustees of the University of Illinois

|   |      |
|---|------|
| accgccatcg gagctctacc cgcaactaatg cctaagtgcac cgccagcaaa atgttgcagc | 1500 |
| actgaccctt ttgggaccgc aatgggttga attagcggaa cgtcgtgtag ggggaaagcg   | 1560 |
| gtcgaccgca ttatcgcttc tccggcggtg gctagcggga agggttgtca atgcgtcgga   | 1620 |
| cttaccgctt accgcgaaac ggaccaaagg ccgtggtctt cgccacggcc tttcgaccga   | 1680 |
| cctcacgcta gaaggactcc ggctatgaca gcagcagggg agtttgcaccg tctacgtgcc  | 1740 |
| aatgctacgc gggtagatgt ggttgcactg gataggtaa tgccagttag gcggcaaaca    | 1800 |
| agggtgcctc ttaggctgcc caacaatgag cgagtgtaaa ttacaactac tttcgaccga   | 1860 |
| tgtccttccg gtctgcgctt aataaaaact accgcaattg agccgcaaag tagacaccac   | 1920 |
| gttgcggcgc acccagccaa tgccggctt gtcagcaaac ggcagactta aactggactc    | 1980 |
| gcgtaaaaat gcgcggcctc ttttggcgga gcgcactac cacgacgcga cctcaactgcc   | 2040 |
| gtcaatagac cttcttagtcc tatacaccgc ctactcgccg taaaaggcac tgcagagcaa  | 2100 |
| cgacgtattt ggctgatgtg ttttagtcgct aaaggtacaa cggtgagcga aattactact  | 2160 |
| aaagtcggcg cgacatgacc tccgacttca agtctacacg ccgctcaacg cactgatgga   | 2220 |
| tgcccattgt caaagaaata ccgtcccact ttgcgtccag cggtcgcccgt ggcgcggaaa  | 2280 |
| gccgcactt taatagctac tcgcaccacc aatacggcta gcgcagtgtg atgcagactt    | 2340 |
| gcagcttttgg ggcttgcaca cctcgcggct ttagggctta gagatagcac gccaccaact  | 2400 |
| tgacgtgtgg cggctgcccgt gcgactaact tcgtcttcgg acgctacagc caaaggcgct  | 2460 |
| ccacgcctaa cttttaccag acgacgacga cttgccgttc ggcaacgact aagctccgca   | 2520 |
| attggcagtgc ctcgttagtag gagacgtacc agtccagttac ctactcgct gctaccacgt | 2580 |
| cctataggac gactacttcg tcttggtaa attgcggcac gcgacaagcg taataggctt    | 2640 |
| ggtaggcgac accatgtgcg acacgctggc gatgccggac atacaccacc tacttcgg     | 2700 |
| ataactttgg gtgccgtacc acggttactt agcagactgg ctactaggcg cgaccgatgg   | 2760 |
| ccgctactcg cttgcgcatt gcgttacca cgtcgctca gcattagttgg gctcacacta    | 2820 |
| gtagaccagc gacccttac ttagtccggt gccgcgatta gtgctgcgcg acatagcgac    | 2880 |
| ctagttaga cagctaggaa gggcgccca cgtcataactt ccgcgcctc ggctgtggtg     | 2940 |
| ccggcttgc taataaacgg gctacatgcg cgccaccta cttctggcgc ggaaggccg      | 3000 |
| acacggctt accaggttgtt ttttaccga aagcgatgga cctctctgcg cggcgacta     | 3060 |
| ggaaacgcctt atgcgggtgc gctaccatt gtcagaaccg ccaaagcgat ttatgaccgt   | 3120 |
| ccgcaaagca gtcatagggg caaatgtccc gccgaagcag accctgaccc acctagtcag   | 3180 |

DRAFT

|   |      |
|---|------|
| cgactaattt atactacttt tgccgttggg caccagccga atgccgccac taaaaccgct   | 3240 |
| atgcggcttg ctagcggtca agacatactt gccagaccag aaacggctgg cgtgcggcgt   | 3300 |
| aggtcgcac tgccttcgtt ttgtggtcgt cgtaaaaag gtcaaggcaa ataggcccgt     | 3360 |
| ttggtagctt cactggtcgc ttatggacaa ggcagtatcg ctattgctcg aggacgtgac   | 3420 |
| ctaccaccgc gacctaccat tcggcgaccg ttcgcccatt cacggagacc tacagcgagg   | 3480 |
| tgttccattt gtcaactaac ttgacggact ttagtggcgtc ggcctctcgc ggcccggttga | 3540 |
| gaccgagtgt catgcgcattc acgttggctt gcgcgtggcgt accagtcttc ggcccggtgt | 3600 |
| gtcgcggacc gtcgtcaccg cagaccgcct tttggagtca cactgcgagg ggcggcgcag   | 3660 |
| ggtgcggtag ggcgttagact ggtggcgtt ttacctaaaa acgtagctcg acccattatt   | 3720 |
| cgcaaccgtt aaattggcggt tcagtccgaa agaaagtgtc tacacctaac cgctatTTT   | 3780 |
| tgttgcgac tgcggcgcacg cgctagtc aa gtggccacag ctatctagac ttgtctttga  | 3840 |
| gtaaaggctt cttctggatc agctggtagt agtagtagta gtggccatTA ttatccatct   | 3900 |
| attcaactgac taatctacgt aaagctgatc tagggagctg gttaaggcca ataaaaggta  | 3960 |
| gtataacggc agaaaaccgt tacactcccg ggcctttgga ccgggacaga agaactgctc   | 4020 |
| gttaaggatcc ccagaaaaggg gagagcgggt tccttacgtt ccagacaact tacagcactt | 4080 |
| ccttcgtcaa ggagaccttc gaagaacttc tgTTTgttgc agacatcgct gggaaacgtc   | 4140 |
| cgtcgccTTT gggggTggac cgctgtccac ggagacgccc gttttcggtg cacatattct   | 4200 |
| atgtggacgt ttccggcgtg ttgggggtcac ggtgcaacac tcaacctatc aacacctttc  | 4260 |
| tcagTTTacc gagaggagtt cgccataagtt gttccccgac ttccTACggg tctttccatgg | 4320 |
| ggtaacatac cctagactag accccggagc cacgtgtacg aaatgtacac aaatcagctc   | 4380 |
| caatTTTTTc cagatccggg gggcttgggt cccctgcacc aaaaggaaac tttttgtgt    | 4440 |
| actattatgg tactTTTtcg gacttgagtg gcgcgtcaga cagctttca aagactagct    | 4500 |
| tttcaagctg tcgcagaggc tggactacgt cgagagcctc ccgcttctta gagcacgaaa   | 4560 |
| gtcgaagcta catcctcccg cacctataca ggacgcccattt atcgacgc ggctaccaaa   | 4620 |
| gatgtttcta gcaataaaaa tagccgtgaa acgtagccgg cgccggggct aaggcTTTca   | 4680 |
| cgaactgtaa ccccttaaat cgcrcTcgga ctggataacg tagagggcgg cacgtgtccc   | 4740 |
| acagtgcac gttctggacg gactttggct tgacggggcga caagacgtcg gccagcgcct   | 4800 |
| ccggTaccTA cgctagcgcac gcccggtaga atcggtctgc tcgcccAAGC cgggtaagcc  | 4860 |
| tggcgTTTCT tagccagtta tggatgtac cgcaactaaag tatacgcgt aacgactagg    | 4920 |

|             |             |             |            |             |            |      |
|-------------|-------------|-------------|------------|-------------|------------|------|
| ggtacacata  | gtgaccgttt  | gacactacct  | gctgtggcag | tcacgcaggc  | agcgcgtccg | 4980 |
| agagctactc  | gactacgaaa  | cccggtcct   | gacggggctt | caggccgtgg  | agcacgtgcg | 5040 |
| cctaaagccg  | aggttgtac   | aggactgcct  | gttaccggcg | tattgtcgcc  | agtaactgac | 5100 |
| ctcgctccgc  | tacaagcccc  | taagggttat  | gctccagcgg | ttgtagaaga  | agacctccgg | 5160 |
| caccaaccga  | acataacctcg | tcgtctgcgc  | gatgaagctc | gcctccgtag  | gcctcgaacg | 5220 |
| tcctagcggc  | gccgaggccc  | gcatatacga  | ggcgtaacca | gaactgcttg  | agatagtctc | 5280 |
| gaaccaactg  | cggttaaagc  | tactacgtcg  | aaccgcgtc  | ccagctacgc  | tgcgttagca | 5340 |
| ggctaggcct  | cggccctgac  | agcccgcatg  | tgttagcgg  | gcgtcttcgc  | gccggcagac | 5400 |
| ctggctaccg  | acacatcttc  | atgagcggct  | atcaccttg  | gctgcggggt  | cgtgagcagg | 5460 |
| ctcccgttcc  | cttatctcat  | ctacggctgg  | ccctagatag | ctattttatt  | ttctaaaata | 5520 |
| aatcagaggt  | ctttttcccc  | ccttactttc  | tgggtggac  | atccaaacgg  | ttcgatcgaa | 5580 |
| ttcattgcgg  | taaaacgttc  | cgtacctttt  | tatgtattga | ctcttatctc  | ttcaagtcta | 5640 |
| gttccagtcc  | ttgtctacct  | tgtcgactta  | tacccggttt | gtcctataga  | caccattcgt | 5700 |
| caaggacggg  | gccgagtcgg  | ggttcttgc   | taccttgcg  | acttatacc   | ggttgcct   | 5760 |
| atagacacca  | ttcgtcaagg  | acggggccga  | gtcccggttc | tgtctacca   | gggtctacg  | 5820 |
| ccaggtcggg  | agtcgtcaaa  | gatctttgg   | tagtctacaa | aggtcccacg  | gggttcctgg | 5880 |
| actttactgg  | gacacgaaat  | aaacttgatt  | ggttagtcaa | gcgaagagcg  | aagacaagcg | 5940 |
| cgcgaagacg  | aggggctcga  | gttattttct  | cgggtgttgg | ggagtgagcc  | ccgcggtcag | 6000 |
| gaggctaact  | gactcagcgg  | gcccatggc   | acataggta  | tttgggagaa  | cgtcaacgta | 6060 |
| ggctgaacac  | cagagcgaca  | aggaaccctc  | ccagaggaga | ctcactaact  | gatggcagt  | 6120 |
| cgcggccaga  | aagtaagtac  | gtcgtacata  | gttttaatta | aaccaaaaaa  | aagaattcat | 6180 |
| aaatgttaatt | taccggtatac | aacgtaatta  | cttagccgt  | tgcgcgcccc  | tctccgccaa | 6240 |
| acgcataacc  | gcgagaaggc  | gaaggagcga  | gtgactgagc | gacgcgagcc  | agcaagccga | 6300 |
| cgcggctcgc  | catagtcgag  | ttagttccg   | ccattatgcc | aataggtgtc  | ttagtcccct | 6360 |
| attgcgtcct  | ttcttgtaca  | ctcggtttcc  | ggtcgtttc  | cggtccttgg  | cattttccg  | 6420 |
| gcgcaacgac  | cgcggggagg  | tatccgagggc | ggggggactg | ctcgttagtgt | tttagctgc  | 6480 |
| gagttcagtc  | tccaccgctt  | tgggtgtcc   | tgtatattct | atggtccgca  | aagggggacc | 6540 |
| ttcgagggag  | cacgcgagag  | gacaaggctg  | ggacggcgaa | tggcctatgg  | acaggcggaa | 6600 |
| agagggaaagc | cttcgcacc   | gcgaaagagt  | atcgagtgcg | acatccata   | agtcaagcc  | 6660 |

catccagcaa gcgagggtcg acccgacaca cgtgcttggg gggcaagtcg ggctggcgac 6720  
 gcggaatagg ccattgatag cagaactca gttggccat tctgtgctga atagcggtga 6780  
 ccgtcgctgg tgaccattgt cctaatcgtc tcgctccata catccgcccac gatgtctcaa 6840  
 gaacttcacc accggattga tgccgatgtg atcttcttgt cataaaccat agacgcgaga 6900  
 cgacttcggt caatggaagc cttttctca accatcgaga actaggccgt ttgtttggtg 6960  
 gcgaccatcg ccaccaaaaa aacaaacgtt cgtcgctaa tgcgcgctt ttttcctag 7020  
 agttcttcta ggaaactaga aaagatgccc cagactgcga gtcaccttgc ttttgagtgc 7080  
 aattccctaa aaccagtact ctaatagttt ttccctagaag tggatctagg aaaatttaat 7140  
 ttttacttca aacgcccccg tttagttaga tttcatatat actcatttga accagactgt 7200  
 caatggttac gaattagtca ctccgtggat agagtcgcta gacagataaa gcaagtaggt 7260  
 atcaacggac tgagggcag cacatctatt gatgctatgc cctccgaat ggtagaccgg 7320  
 ggtcacgacg ttactatggc gctctgggtg cgagtggccg aggtctaaat agtcgttatt 7380  
 tggtcggtcg gccttcccg ctcgcgtctt caccaggacg ttgaaatagg cggaggtagg 7440  
 tcagataatt aacaacggcc ctgcgtctc attcatcaag cggtaatttcaaaacgcgt 7500  
 tgcaacaacg gtaacgatgt ccgttagcacc acagtgcgag cagcaaacca taccgaagta 7560  
 agtcgaggcc aagggttgct agttccgctc aatgtactag ggggtacaac acgtttttc 7620  
 gccaatcgag gaagccagga ggctagcaac agtcttcatt caaccggcgt cacaatagtg 7680  
 agtaccaata ccgtcgac gtatagatg aatgacagta cggtaggcatt tctacgaaaa 7740  
 gacactgacc actcatgagt tggtttagta agactcttac cacatacgcc gctggctcaa 7800  
 cgagaacggg ccgcagttat gccctattat ggcgcgggtgt atcgtcttga aattttcacg 7860  
 agtagtaacc ttttgcaga agccccgctt ttgagagttc cttagaatggc gacaactcta 7920  
 ggtcaagcta cattgggtga gcacgtgggt tgactagaag tcgtagaaaa tgaaagtgg 7980  
 cgcaaagacc cactcgaaaaacttac ttttacggcg tttttccct tattcccgct 8040  
 gtgcctttac aacttatgag tatgagaagg aaaaagttt aataacttcg taaatagtcc 8100  
 caataacaga gtactcgccct atgtataaaac ttacataaaat ctttttatgg 8160  
 aaggcgcgtg taaag 8175

<210> 8  
 <211> 8161  
 <212> DNA  
 <213> Unknown

<220>  
<223> pICAST OMN.

<400> 8

|                        |   |      |
|------------------------|---|------|
| ctgcagcctg aatatgggcc  | aaacaggata tctgtggtaa gcagttcctg ccccggtca      | 60   |
| ggccaagaa cagatggaac   | agctgaatat gggccaaaca ggatatctgt ggtaaggcgt     | 120  |
| tcctgccccg gctcagggcc  | aagaacagat ggtccccaga tgccgtccag ccctcagcag     | 180  |
| tttctagaga accatcagat  | gttccaggg tgcccaagg acctgaaatg accctgtgcc       | 240  |
| ttatgtAAC taaccaatca   | gttcgcttct cgcttctgtt cgccgccttc tgctccccga     | 300  |
| gctcaataaa agagccaca   | accctcact cggggcggca gtcctccgat tgactgagtc      | 360  |
| gcccggtac ccgtgtatcc   | aataaaccct cttgcagttt catccgactt gtggctcgc      | 420  |
| tgttccttgg gagggtctcc  | tctgagtgtat tgactacccg tcaggggggg tcttcattt     | 480  |
| gggggctcgt ccgggatcgg  | gagacccctg cccagggacc accgaccac caccggagg       | 540  |
| caagctggcc agcaacttat  | ctgtgtctgt ccgattgtct agtgtctatg actgatttt      | 600  |
| tgcgcctgcg tcggtaact   | ttagctaact agctctgtat ctggcggacc cgtggtgaa      | 660  |
| ctgacgagtt ctgaacaccc  | ggccgcaacc ctggagacg tcccaggac tttgggggcc       | 720  |
| gtttttgtgg cccgaccta   | ggaaggaggt cgatgtggaa tccgaccccg tcaggatatg     | 780  |
| tggttcttgtt aggagacgag | aacctaacc acgtccgc tccgtctgaa ttttgcttt         | 840  |
| cgttttggaa ccgaagccgc  | gcgtcttgc tgctgcagca tcgttctgtt ttgtctctgt      | 900  |
| ctgactgtgt ttctgtat    | ttt gtctgaaaat tagggccaga ctgttaccac tcccttaagt | 960  |
| ttgaccttag gtaactggaa  | agatgtcgag cggctcgctc acaaccagtc ggtagatgtc     | 1020 |
| aagaagagac gttgggttac  | cttctgctct gcagaatggc caaccttaa cgtcgatgg       | 1080 |
| cccgagacg gcaccttaa    | ccgagacctc atcaccagg ttaagatcaa ggtctttca       | 1140 |
| cctggccgc atggacaccc   | agaccaggtc ccctacatcg tgacctggaa agccttggct     | 1200 |
| tttgcacccct ctccctgggt | caagcccttt gtacacccta agcctccgcc tcctttcct      | 1260 |
| ccatccgccc cgtctctccc  | ccttgaacct ctcgttcga cccgcctcg atcctccctt       | 1320 |
| tatccagccc tcactccttc  | tctaggcgcc ggccgctcta gcccattaat acgactact      | 1380 |
| atagggcgat tcgaacacca  | tgacccatca tcacatcac gtcgacgaac agaaactcat      | 1440 |
| ttccgaagaa gacctactcg  | agatgggcgt gattacggat tcaactggccg tcgtttaca     | 1500 |
| acgtcggtac tggaaaacc   | ctggcggtac ccaacttaat cgccttgcaag cacatcccc     | 1560 |
| tttcgcccagc tggcgtaata | gcaagaggc cgcaccgat cgccttccc aacagttacg        | 1620 |

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| cagcctgaat | ggcgaatggc | gctttgcctg | gttccggca  | ccagaagcgg | tgccggaaag | 1680 |
| ctggctggag | tgcgatcttc | ctgaggccga | tactgtcgtc | gtcccctcaa | actggcagat | 1740 |
| gcacggttac | gatgcgcccc | tctacaccaa | cgtgacctat | cccattacgg | tcaatccgcc | 1800 |
| gttgttccc  | acggagaatc | cgacgggttg | ttactcgctc | acatttaatg | ttgatgaaag | 1860 |
| ctggctacag | gaaggccaga | cgcgaattat | ttttgatggc | gttaactcgg | cgtttcatct | 1920 |
| gtggtgcaac | ggcgctggg  | tcggttacgg | ccaggacagt | cgtttgccgt | ctgaatttga | 1980 |
| cctgagcgca | tttttacgcg | ccggagaaaa | ccgcctcgcg | gtgatggtgc | tgcgctggag | 2040 |
| tgacggcagt | tatctggaag | atcaggatat | gtggcggatg | agcggcattt | tccgtgacgt | 2100 |
| ctcgttgctg | cataaaccga | ctacacaaat | cagcgatttc | catgttgcca | ctcgctttaa | 2160 |
| tgatgatttc | agccgcgctg | tactggaggc | tgaagttcag | atgtgcggcg | agttgcgtga | 2220 |
| ctacctacgg | gtaacagttt | ctttatggca | gggtgaaacg | caggtcgcca | gcggcaccgc | 2280 |
| gccttcggc  | ggtgaaatta | tcgatgagcg | tggtggttat | gccgatcgcg | tcacactacg | 2340 |
| tctgaacgtc | gaaaacccga | aactgtggag | cgccgaaatc | ccgaatctct | atcgtgcgg  | 2400 |
| ggttgaactg | cacaccgccc | acggcacgct | gattgaagca | gaagcctgcg | atgtcggtt  | 2460 |
| ccgcgagggt | cggattgaaa | atggtctgct | gctgctgaac | ggcaagccgt | tgctgattcg | 2520 |
| aggcgtaac  | cgtcacgagc | atcatcctct | gcatggtcag | gtcatggatg | agcagacgat | 2580 |
| ggtgcaggat | atcctgctga | tgaagcagaa | caacttaac  | gccgtcgct  | gttcgcatta | 2640 |
| tccgaaccat | ccgctgtggt | acacgctgtg | cgaccgctac | ggcctgtatg | tggtggatga | 2700 |
| agccaatatt | gaaaccacg  | gcatggtgcc | aatgaatcgt | ctgaccgatg | atccgcgtg  | 2760 |
| gctaccggcg | atgagcgaac | gcgtaacgcg | aatggtgcag | cgcgatcgta | atcaccggag | 2820 |
| tgtgatcatc | tggtcgttgg | ggaatgaatc | aggccacggc | gctaattcag | acgcgctgta | 2880 |
| tcgctggatc | aaatctgtcg | atccctcccg | cccggtgca  | tatgaaggcg | gcggagccga | 2940 |
| caccacggcc | accgatatta | tttggccgat | gtacgcgcgc | gtggatgaag | accagccctt | 3000 |
| cccgctgtg  | ccgaaatgg  | ccatcaaaaa | atggcttcg  | ctacctggag | agacgcgccc | 3060 |
| gctgatcctt | tgcgaatacg | cccacgcgt  | ggtaacagt  | cttggcggtt | tcgctaaata | 3120 |
| ctggcaggcg | tttcgtcagt | atccccgtt  | acagggcggc | ttcgtctggg | actgggtgga | 3180 |
| tcagtcgctg | attaaatatg | ataaaaacgg | caacccgtgg | tcggcttacg | gcggtgattt | 3240 |
| tggcgatacg | ccgaacgatc | gccagttctg | tatgaacggt | ctggctttg  | ccgaccgcac | 3300 |
| gccgcatcca | gcgctgacgg | aagcaaaaca | ccagcagcag | tttttccagt | tccgtttatc | 3360 |

cgggcaaac atcgaagtga ccagcgaata cctgtccgt catagcgata acgagctcct 3420  
gcactggatg gtggcgctgg atggtaagcc gctggcaagc ggtgaagtgc ctctggatgt 3480  
cgctccacaa ggttaaacagt tgattgaact gcctgaacta ccgcagccgg agagcgccgg 3540  
gcaactctgg ctcacagtac gcgttagtgca accgaacgcg accgcattgtt cagaagccgg 3600  
gcacatcagc gcctggcagc agtggcgtct ggccggaaac ctcagtgta cgctccccgc 3660  
cgcgccccac gccatcccgc atctgaccac cagcgaaatg gattttgca tcgagctggg 3720  
taataagcgt tggcaattta accgccagtc aggcttctt tcacagatgt ggattggcga 3780  
taaaaaacaa ctgctgacgc cgctgcgcga tcagttcacc cgtgtcgata gatctggagg 3840  
tggcggcgc aggcccttgc gcgcggatc cttaaattaa aattgaccgg taataatagg 3900  
tagataagtg actgattaga tgcatttcga cttagatccct cgaccaattc cggttatttt 3960  
ccaccatatt gccgtctttt ggcaatgtga gggcccgaa acctggccct gtcttcttga 4020  
cgagcattcc taggggtctt tccccctctcg ccaaaggaat gcaaggtctg ttgaatgtcg 4080  
tgaaggaagc agttcctctg gaagcttctt gaagacaaac aacgtctgtt ggcgacccttt 4140  
gcaggcagcg gaaccccccga cctggcgaca ggtgcctctg cggccaaaag ccacgtgtat 4200  
aagatacacc tccaaaggcg gcacaacccc agtgcacgt tgtgagttgg atagttgtgg 4260  
aaagagtcaa atggctctcc tcaagcgtat tcaacaaggg gctgaaggat gcccagaagg 4320  
tacccattt tatggatct gatctggggc ctcggtgac atgcatttaca tgtgttttagt 4380  
cgaggttaaa aaacgtctag gccccccgaa ccacggggac gtggtttcc tttgaaaaac 4440  
acgatgataa taccatgaaa aagcctgaac tcaccgcac gtctgtcgag aagtttctga 4500  
tcgaaaagtt cgacagcgtc tccgacactga tgcagctctc ggagggcgaa gaatctcg 4560  
ctttcagctt cgatgttagga gggcgtggat atgtcctgcg ggttaatagc tgcgcccgt 4620  
gtttctacaa agatcgttat gtttatcggc actttgcac ggccgcgcgc ccgattccgg 4680  
aagtgcgttga cattggggaa tttagcgaga gcctgaccta ttgcatttcc cggccgtgcac 4740  
agggtgtcac gttgcaagac ctgcctgaaa ccgaactgcc cgctgttctg cagccggcgt 4800  
cgaggccat ggtgcgatc gctggggccg atcttagcca gacgagcggg ttcggcccat 4860  
tcggaccgcgca aggaatcggt caatacacta catggcgtga tttcatatgc gcgattgtcg 4920  
atccccatgt gtatcactgg caaactgtga tggacgacac cgtagtgcg tccgtcgccgc 4980  
aggctctcgatc tgagctgtatc ctttggggccg aggactgcgc cgaagtccgg cacctcgatc 5040  
acqcgqgatattt cggctccaaac aatgtcctgaa cggacaatgg ccgcataaca gcggtcattg 5100

actggagcga ggcgatgttc ggggattccc aatacgaggt cgccaacatc ttcttctgga 5160  
 ggccgtggtt ggcttgtatg gagcagcaga cgcgctactt cgagcggagg catccggagc 5220  
 ttgcaggatc gccgcggctc cgggcgtata tgctccgcat tggtcttgac caactctatc 5280  
 agagcttggt tgacggcaat ttcgatgtatc cagcttggc gcagggtcga tgcgacgcaa 5340  
 tcgtccgatc cggagccggg actgtcgggc gtacacaaat cgcccgaga agcgcggccg 5400  
 tctggaccga tggctgtgtt gaagtactcg ccgatagtgg aaaccgacgc cccagcactc 5460  
 gtccgagggc aaaggaatag agtagatgcc gaccgggatc tatcgataaa ataaaagatt 5520  
 ttattnatc tccagaaaaa ggggggaatg aagacccaa cctgttagtt tggcaagcta 5580  
 gcttaagtta cgccattttcaaggcatgg aaaaatacat aactgagaat agagaagttc 5640  
 agatcaaggt caggaacaga tggAACAGCT gaatatggc caaacaggat atctgtggta 5700  
 agcagttcct gccccggctc agggccaaga acagatggaa cagctgaata tggccaaac 5760  
 aggatatctg tggtaagcag ttccctgcccc ggctcaggc caagaacaga tggtccccag 5820  
 atgcggtcca gcccctcagca gtttctagag aaccatcaga tggcccccagg gtgccccaaag 5880  
 gacctgaaat gaccctgtgc ctatTTGAA ctaaccaatc agttcgcttc tcgcttctgt 5940  
 tcgcgcgtt ctgctccccg agctcaataa aagagccac aaccctcac tcggggcgcc 6000  
 agtcctccga ttgactgagt cggccggta cccgtgtatc caataaacc tcttgcagtt 6060  
 gcatccgact tgggtctcg ctgttccttggaggggtctc ctctgagtga ttgactaccc 6120  
 gtcagcgggg gtcttcatt catgcagcat gtatcaaaat taattgggtt tttttctta 6180  
 agtatttaca ttaaatggcc atagttcat taatgaatcg gccaacgcgc ggggagaggc 6240  
 ggTTTGCgtt ttggcgtct tccgcttcct cgctcaactga ctgcgtgcgc tcggcgttc 6300  
 ggctgcggcg agcggatca gctcaactaa aggccgtat acggttatcc acagaatcag 6360  
 gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa 6420  
 aggccgcgtt gctggcgTTT ttccataggc tccgcgggggg tgacgagcat cacaaaaatc 6480  
 gacgctcaag tcagagggtgg cggaaacccga caggactata aagataaccag gcgttcccc 6540  
 ctggaaagctc cctcgtgcgc tccctgttc cgaccctgcgctt tacctgtccg 6600  
 cctttctccc ttggaaagc gtggcgctt ctcatacgatc acgctgttagg tatctcagtt 6660  
 cgggttaggt cggtcgctcc aagctggct gtgtgcacga acccccccgtt cagcccgacc 6720  
 gctgcgcctt atccggtaac tatcgtcttgg agtccaaccc ggtaagacac gacttatacg 6780  
 cactggcagc agccactggt aacaggatta gcagagcggag gtatgttaggc ggtgctacag 6840

|             |            |            |             |              |             |      |
|-------------|------------|------------|-------------|--------------|-------------|------|
| agttcttcaa  | gtggtggcct | aactacggct | acactagaag  | aacagtattt   | ggtatctgcg  | 6900 |
| ctctgctgaa  | gccagttacc | ttcggaaaaa | gagttggtag  | ctcttgcattcc | ggcaaacaaa  | 6960 |
| ccaccgctgg  | tagcggtgg  | tttttggttt | gcaaggcagca | gattacgcgc   | agaaaaaaaag | 7020 |
| gatctcaaga  | agatcctttg | atctttctca | cggggctgt   | cgctcagtg    | aacgaaaact  | 7080 |
| cacgttaagg  | gattttggtc | atgagattat | caaaaaggat  | cttcacccat   | atcctttgc   | 7140 |
| ggccgcaat   | caatctaaag | tatatatgag | taaacttggt  | ctgacagtt    | ccaatgctta  | 7200 |
| atcagtgagg  | cacctatctc | agcgatctgt | ctattcggt   | catccatagt   | tgcctgactc  | 7260 |
| cccgctgtgt  | agataactac | gatacgggag | ggcttaccat  | ctggcccccag  | tgctgcaatg  | 7320 |
| ataccgcgag  | acccacgctc | accggctcca | gatttatcag  | caataaacca   | gccagccgga  | 7380 |
| agggccgagc  | gcagaagtgg | tcctgcaact | ttatccgcct  | ccatccagtc   | tattaattgt  | 7440 |
| tgccggaaag  | ctagagtaag | tagttcgcca | gttaatagtt  | tgcgcaacgt   | tgttgccatt  | 7500 |
| gctacaggca  | tcgtgggtgc | acgctcggt  | tttggtatgg  | cttcattcag   | ctccgggttcc | 7560 |
| caacgatcaa  | ggcgagttac | atgatcccc  | atgttgc     | aaaaagcggt   | tagctccccc  | 7620 |
| ggtcctccga  | tcgttgc    | aagtaagttg | gccgcagtgt  | tatcactcat   | ggttatggca  | 7680 |
| gcactgcata  | attctcttac | tgtcatgcca | tccgttaagat | gctttctgt    | gactggtag   | 7740 |
| tactcaacca  | agtcattctg | agaatagtgt | atgcggcgcac | cgagttgc     | ttgcccggcgc | 7800 |
| tcaatacggg  | ataataccgc | gccacatagc | agaactttaa  | aagtgc       | cattggaaaa  | 7860 |
| cgttcttcgg  | ggcgaaaact | ctcaaggatc | ttaccgctgt  | tgagatccag   | ttcgatgtaa  | 7920 |
| cccaactcg   | caccaactg  | atcttcagca | tctttactt   | tcaccagcgt   | ttctgggtga  | 7980 |
| gcaaaaacag  | gaaggcaaaa | tgccgcaaaa | aagggataaa  | gggcgacacg   | gaaatgttga  | 8040 |
| ataactcatac | tcttcctttt | tcaatattat | tgaagcattt  | atcagggtta   | ttgtctcatg  | 8100 |
| agcggataca  | tatttgaatg | tattnagaaa | aataaacaaa  | taggggttcc   | gcccacattt  | 8160 |
| c           |            |            |             |              |             | 8161 |

<210> 9  
 <211> 8161  
 <212> DNA  
 <213> Unknown

<220>  
 <223> pICAST OMN.

<400> 9  
 gacgtcgac ttataccgg tttgtcctat agacaccatt cgtcaaggac gggcccgagt 60

cccggttctt gtctaccttg tcgacttata cccggttgt cctatacgaca ccattcgta 120  
 aggacggggc cgagtcggc ttcttgccta ccaggggtct acgcccaggc gggagtcgtc 180  
 aaagatctct tggtagtcta caaaggccc acggggttcc tggactttac tggacacgg 240  
 aataaaacttg attggtagt caagcgaaga gcgaagacaa gcgcgcgaag acgagggct 300  
 cgagttattt tctcggtgt tggggagtga gcccccggt caggaggcta actgactcag 360  
 cgggcccattt ggcacatagg ttatttggga gaacgtcaac gtggctgaa caccagagcg 420  
 acaaggaacc ctcccagagg agactcacta actgatggc agtcgcccc agaaagtaaa 480  
 ccccccggca ggccttagcc ctctggggac gggccctgg tggctgggt gtggccctcc 540  
 gttcgaccgg tcgttgaata gacacagaca ggctaacaga tcacagatac tgactaaaat 600  
 acgcggacgc agccatgatc aatcgattga tcgagacata gaccgcctgg gcaccacctt 660  
 gactgctcaa gacttgggg ccggcggttgg gaccctctgc agggccctg aaaccccccgg 720  
 caaaaacacc gggctggact cttccctca gctacacctt aggctgggc agtcctatac 780  
 accaagacca tcctctgctc ttggattttt tcaagggcgg aggccagactt aaaaacgaaa 840  
 gccaaacctt ggcttcggcg cgccagaacag acgacgtcgt agcaagacac aacagagaca 900  
 gactgacaca aagacataaa cagactttt atcccggtct gacaatggt agggattca 960  
 aactggaatc cattgacctt tctacagctc gccgagcggag tgggttgcag ccatctacag 1020  
 ttcttctctg caacccaaatg gaagacgaga cgtcttaccc gttggaaatt gcagcctacc 1080  
 ggcgtctgc cgtggaaatt ggctctggag tagtgggtcc aattctagtt ccagaaaagt 1140  
 ggaccgggccc tacctgtggg tctggtccag gggatgttagc actggaccc tcggaaaccga 1200  
 aaactggggg gagggaccca gttcggaaa catgtggat tcggaggcgg aggagaagga 1260  
 ggtaggcggg gcagagaggg ggaacttggg ggagcaagct gggccggagc taggaggaa 1320  
 ataggtcggg agtgaggaag agatccgcgg ccggcgagat cggtaatta tgctgagtga 1380  
 tatcccgcta agcttgggt acgtggtagt agtagtagt cagctgcttgc tctttgagta 1440  
 aaggcttctt ctggatgagc tctacccgca ctaatgccta agtgcacggc agcaaaatgt 1500  
 tgcagcactg accctttgg gaccgcaatg ggtgaatta gcggAACgtc gtgtaggggg 1560  
 aaagcggtcg accgcattat cgcttctccg ggcgtggcta gcgggaagggt ttgtcaatgc 1620  
 gtcggactta cgccttaccc cgaaacggac caaaggccgt ggtcttcggcc acggcccttc 1680  
 gaccgaccc acgctagaag gactccggct atgacagcag cagggaggt tgaccgtcta 1740  
 cgtgccaatg ctacgcgggt agatgtgggt gcactggata ggttaatgcc agttaggcgg 1800

DRAFT

|             |            |            |             |             |            |      |
|-------------|------------|------------|-------------|-------------|------------|------|
| caaacaaggg  | tgcctcttag | gctgcccaac | aatgagcgag  | tgtaaattac  | aactacttc  | 1860 |
| gaccgatgtc  | cttccggct  | gcgcttaata | aaaactaccg  | caattgagcc  | gcaaagtaga | 1920 |
| caccacgtt   | cccgcgaccc | agccaatgcc | ggtcctgtca  | gcaaacggca  | gacttaaact | 1980 |
| ggactcgct   | aaaaatgcgc | ggcctttt   | ggcggagcgc  | cactaccacg  | acgcgacctc | 2040 |
| actgcccgtca | atagaccttc | tagtcctata | caccgcctac  | tcgcccgtaaa | aggcactgca | 2100 |
| gagcaacgac  | gtatggct   | gatgtgttta | gtcgctaaag  | gtacaacggt  | gagcgaaatt | 2160 |
| actactaaag  | tcggcgac   | atgacctccg | acttcaagtc  | tacacgccgc  | tcaacgcact | 2220 |
| gatggatgcc  | cattgtcaaa | gaaataccgt | cccacttgc   | gtccagcggt  | cgcgtggcg  | 2280 |
| cgaaaaagccg | ccacttaat  | agctactcgc | accaccaata  | cggttagcgc  | agtgtgatgc | 2340 |
| agacttgcag  | cttttggct  | ttgacacctc | gcggctttag  | ggcttagaga  | tagcacgcca | 2400 |
| ccaacttgcac | gtgtggcg   | tgccgtgcga | ctaacttcgt  | cttcggacgc  | tacagccaaa | 2460 |
| ggcgctccac  | gcctaactt  | taccagacga | cgacgacttg  | ccgttcggca  | acgactaagc | 2520 |
| tccgcaattt  | gcagtgtcg  | tagtaggaga | cgtaccagtc  | cagtacctac  | tcgtctgcta | 2580 |
| ccacgtccta  | taggacgact | acttcgtctt | gttggaaattt | cggcacgcga  | caagcgtaat | 2640 |
| aggcttggta  | ggcgacacca | tgtgcgacac | gctggcgatg  | ccggacatac  | accacctact | 2700 |
| tcggttataa  | ctttgggtgc | cgtaccacgg | ttacttagca  | gactggctac  | taggcgcgac | 2760 |
| cgatggccgc  | tactcgctt  | cgcattgcgc | ttaccacgtc  | gcgctagcat  | tagtggctc  | 2820 |
| acactagtag  | accagcgacc | ccttacttag | tccggtgccg  | cgattagtgc  | tgcgcgacat | 2880 |
| agcgacctag  | tttagacagc | taggaagggc | ggccacgtc   | atactccgc   | cgcctcggt  | 2940 |
| gtgggtccgg  | tggctataat | aaacgggcta | catgcgcg    | cacctacttc  | tggtcggaa  | 3000 |
| ggccgcacac  | ggcttacca  | ggtagtttt  | taccgaaagc  | gatggacctc  | tctgcgcgg  | 3060 |
| cgacttagaa  | acgcttatgc | gggtgcgcta | cccattgtca  | gaaccgcacaa | agcgatttat | 3120 |
| gaccgtccgc  | aaagcagtca | tagggcaaa  | tgtccgcgc   | aagcagaccc  | tgacccacct | 3180 |
| agtcagcgac  | taatttatac | tactttgcc  | gttgggcacc  | agccgaatgc  | cgcactaaa  | 3240 |
| accgctatgc  | ggcttgctag | cggtaagac  | atacttgcca  | gaccagaaac  | ggctggcg   | 3300 |
| cggcgttaggt | cgcgactgcc | ttcggtttgt | ggtcgtcg    | aaaaaggtca  | aggcaaata  | 3360 |
| gcccggtttgg | tagcttact  | ggtcgtttat | ggacaaggca  | gtatcgctat  | tgctcgagga | 3420 |
| cgtgacctac  | caccgcgacc | taccattcgg | cgaccgttcg  | ccacttcacg  | gagacctaca | 3480 |
| gcgagggttt  | ccatttgtca | actaacttga | cgacttgat   | ggcgtcg     | ccgtcg     | 3540 |

|            |             |            |            |             |             |      |
|------------|-------------|------------|------------|-------------|-------------|------|
| cgttgagacc | gagtgtcatg  | cgcacacgt  | tggcttgcgc | tggcgtacca  | gtcttcggcc  | 3600 |
| cgtgtagtcg | cggaccgtcg  | tcaccgcaga | ccgcctttg  | gagtcacact  | gcgaggggcg  | 3660 |
| gcmcagggtg | cggtagggcg  | tagactggtg | gtcgcttac  | ctaaaaacgt  | agctcgaccc  | 3720 |
| attattcgca | accgttaat   | tggcggtcag | tccgaaagaa | agtgtctaca  | cctaaccgct  | 3780 |
| atttttgtt  | gacgactgcg  | gcmcagcgct | agtcaagtgg | gcacagctat  | ctagacctcc  | 3840 |
| accaccgtcg | tccggaaccg  | cgcggctag  | gaattaattg | ttaactggcc  | attattatcc  | 3900 |
| atctattcac | tgactaatct  | acgtaaagct | gatctaggga | gctggtaag   | gccaaataaaa | 3960 |
| ggtgtataa  | cggcagaaaa  | ccgttacact | cccgccctt  | tggaccggga  | cagaagaact  | 4020 |
| gctcgtaagg | atccccagaa  | aggggagagc | ggttcctta  | cgttccagac  | aacttacagc  | 4080 |
| acttccttcg | tcaaggagac  | cttcgaagaa | cttctgtttg | ttgcagacat  | cgctggaaa   | 4140 |
| cgtccgtcgc | cttgggggt   | ggaccgctgt | ccacggagac | gccggtttc   | ggtcacata   | 4200 |
| ttctatgtgg | acgttccgc   | cgtgttgggg | tcacggtgca | acactcaacc  | tatcaacacc  | 4260 |
| tttctcagtt | taccgagagg  | agttcgcata | agttgtccc  | cgacttccta  | cgggtcttcc  | 4320 |
| atggggtaac | ataccctaga  | ctagaccccg | gagccacgtg | tacgaaatgt  | acacaaatca  | 4380 |
| gctccaattt | tttgcagatc  | cggggggctt | ggtgcccctg | cacccaaagg  | aaacttttg   | 4440 |
| tgctactatt | atggtacttt  | ttcggacttg | agtggcgctg | cagacagctc  | ttcaaagact  | 4500 |
| agctttcaa  | gctgtcgca   | aggctggact | acgtcgagag | cctcccgctt  | cttagagcac  | 4560 |
| gaaagtgcga | gctacatcct  | cccgcaccta | tacaggacgc | ccatttatcg  | acgcggctac  | 4620 |
| caaagatgtt | tctagcaata  | caaatacgcc | tgaaacgtag | ccggcgcgag  | ggctaaggcc  | 4680 |
| ttcacgaact | gtaaccctt   | aaatcgctc  | cggactggat | aacgtagagg  | gcggcacgtg  | 4740 |
| tcccacagt  | caacgttctg  | gacggacttt | ggcttgcgg  | gcgacaagac  | gtcggccagc  | 4800 |
| gcctccggta | cctacgctag  | cgacgccggc | tagaatcggt | ctgctcgccc  | aagccggta   | 4860 |
| agcctggcgt | tccttagcca  | gttatgtgat | gtaccgcact | aaagtatacg  | cgctaacgac  | 4920 |
| taggggtaca | catagtgacc  | gtttgacact | acctgctgtg | gcagtcacgc  | aggcagcgcg  | 4980 |
| tccgagagct | actcgactac  | gaaacccggc | tcctgacggg | gcttcaggcc  | gtggagcacg  | 5040 |
| tgcgcctaaa | gccgaggttg  | ttacaggact | gcctgttacc | ggcgtattgt  | cgccagtaac  | 5100 |
| tgacctcgct | ccgctacaag  | cccctaaggg | ttatgctcca | gcgggtgttag | aagaagacct  | 5160 |
| ccggcaccaa | ccgaacatac  | ctcgtcgtct | gcgcgtgaa  | gctcgctcc   | gtaggcctcg  | 5220 |
| aacgtcctag | cggcggcggag | gccccgtat  | acgaggcgta | accagaactg  | gttgagatag  | 5280 |

|             |             |             |            |             |             |      |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| tctcgaacca  | actgccgtta  | aagctactac  | gtcgaacccg | cgtcccagct  | acgctgcgtt  | 5340 |
| agcaggctag  | gcctcgcccc  | tgacagcccg  | catgtttta  | gcgggcgtct  | tcgcgcggc   | 5400 |
| agacctggct  | accgacacat  | cttcatgagc  | ggctatcacc | tttggctgag  | gggtcgtgag  | 5460 |
| caggctcccg  | tttccttatac | tcatctacgg  | ctggccctag | atagctattt  | tattttctaa  | 5520 |
| aataaaatcag | aggtcttttt  | ccccccttac  | tttctggggt | ggacatccaa  | accgttcgat  | 5580 |
| cgaattcatt  | gcggtaaaac  | gttccgtacc  | tttttatgta | ttgactctta  | tctcttcaag  | 5640 |
| tctagttcca  | gtccttgcgt  | accttgcga   | cttatacccg | gtttgtccta  | tagacaccat  | 5700 |
| tcgtcaagga  | cggggccgag  | tcccggttct  | tgtctacctt | gtcgacttat  | acccggtttg  | 5760 |
| tcctatagac  | accattcgtc  | aaggacgggg  | ccgagtcgg  | gttcttgcgt  | accaggggtc  | 5820 |
| tacgccaggt  | cgggagtcgt  | caaagatctc  | ttggtagtct | acaaagggtcc | cacggggttc  | 5880 |
| ctggacttta  | ctgggacacg  | gaataaactt  | gattggtag  | tcaagcgaag  | agcgaagaca  | 5940 |
| agcgcgcgaa  | gacgaggggc  | tcgagttatt  | ttctcgggtg | ttggggagtg  | agccccgcgg  | 6000 |
| tcaggaggct  | aactgactca  | gcgggcccatt | ggcacatag  | gttattttggg | agaacgtcaa  | 6060 |
| cgtaggctga  | acaccagagc  | gacaaggaac  | cctccagag  | gagactca    | aactgtatggg | 6120 |
| cagtcgcccc  | cagaaagtaa  | gtacgtcgta  | catagttta  | attaaaccaa  | aaaaaagaat  | 6180 |
| tcataaaatgt | aatttacccg  | tatcaacgta  | attacttagc | cggttgcgcg  | ccctctccg   | 6240 |
| ccaaacgcatt | aaccgcgaga  | aggcgaagga  | gcgagtgact | gagcgcacgcg | agccagcaag  | 6300 |
| ccgacgcccgc | tcgccatagt  | cgagttagtt  | tccgcattta | tgccaatagg  | tgtcttagtc  | 6360 |
| cccttattgct | tcctttcttg  | tacactcggt  | ttccgggtcg | tttccggtcc  | ttggcatttt  | 6420 |
| tccggcgcaa  | cgaccgcaaa  | aaggtatccg  | aggcgggggg | actgctcgta  | gtgttttag   | 6480 |
| ctgcgagttc  | agtctccacc  | gctttggct   | gtcctgatat | ttctatggtc  | cgcaaagggg  | 6540 |
| gaccattcgag | ggagcacgcg  | agaggacaag  | gctggacgg  | cgaatggcct  | atggacaggc  | 6600 |
| ggaaagaggg  | aagccattcg  | caccgcgaaa  | gagtagtgcg | tgcgacatcc  | atagagtcaa  | 6660 |
| gccacatcca  | gcaagcgagg  | ttcgacccga  | cacacgtgt  | tggggggcaa  | gtcgggctgg  | 6720 |
| cgacgcggaa  | taggccattt  | atagcagaac  | tcaggttggg | ccattctgtg  | ctgaatagcg  | 6780 |
| gtgaccgtcg  | tcggtgacca  | ttgtccta    | cgtctcgctc | catacatccg  | ccacgatgtc  | 6840 |
| tcaagaactt  | caccaccgga  | ttgatgccga  | tgtgatcttc | ttgtcataaaa | ccatagacgc  | 6900 |
| gagacgactt  | cggtcaatgg  | aagcctttt   | ctcaaccatc | gagaactagg  | ccgtttgttt  | 6960 |
| ggtggcgacc  | atcgccacca  | aaaaaaca    | cggtcgatgt | ctaattgcgcg | tcttttttc   | 7020 |

|   |      |
|---|------|
| ctagagttct tctaggaaac tagaaaagat gccccagact gcgagtcacc ttgttttga    | 7080 |
| gtgcaattcc ctaaaaccag tactctaata gttttccta gaagtggatc tagaaaacg     | 7140 |
| ccggcggtta gttagatttc atatatactc atttgaacca gactgtcaat ggtagtacat   | 7200 |
| tagtcactcc gtggatagag tcgctagaca gataaagcaa gtaggtatca acggactgag   | 7260 |
| gggcagcaca tctattgatg ctatgcctc ccgaatggta gaccggggtc acgacgttac    | 7320 |
| tatggcgctc tgggtgcgag tggccgaggt ctaaatagtc gttatttggt cggtcggcct   | 7380 |
| tcccggtcg cgtcttcacc aggacgttga aataggcga ggttaggtcag ataattaaca    | 7440 |
| acggcccttc gatctcattc atcaagcggt caattatcaa acgcgttgca acaacgtaa    | 7500 |
| cgtatgtccgt agcaccacag tgcgagcagc aaaccatacc gaagtaagtc gaggccaagg  | 7560 |
| gttgcttagtt ccgctcaatg tactaggggg tacaacacgt ttttcgcca atcgaggaag   | 7620 |
| ccaggaggct agcaacagtc ttcatcaac cggcgtcaca atagtgagta ccaataccgt    | 7680 |
| cgtgacgtat taagagaatg acagtacggt aggattcta cgaaaagaca ctgaccactc    | 7740 |
| atgagtttgt tcagtaagac tcttatacaca tacgcccgtg gctcaacgag aacgggcccgc | 7800 |
| agttatgccc tattatggcg cggtgtatcg tcttggaaatt ttcacgagta gtaaccttt   | 7860 |
| gcaagaagcc ccgctttga gagttccctag aatggcgaca actctaggc aagctacatt    | 7920 |
| gggtgagcac gtgggttgac tagaagtcgt agaaaatgaa agtggtcgca aagaccact    | 7980 |
| cgtttttgtc cttccgtttt acggcgttt ttcccttatt cccgctgtgc ctttacaact    | 8040 |
| tatgagtagt agaaggaaaa agttataata acttcgtaaa tagtcccaat aacagagtac   | 8100 |
| tcgcctatgt ataaacttac ataaatctt ttatttgggtt atccccaaagg cgctgtaaa   | 8160 |
| g   | 8161 |